Microarrays

New England Bioterrorism Preparedness Workshop

Dr. David Walt

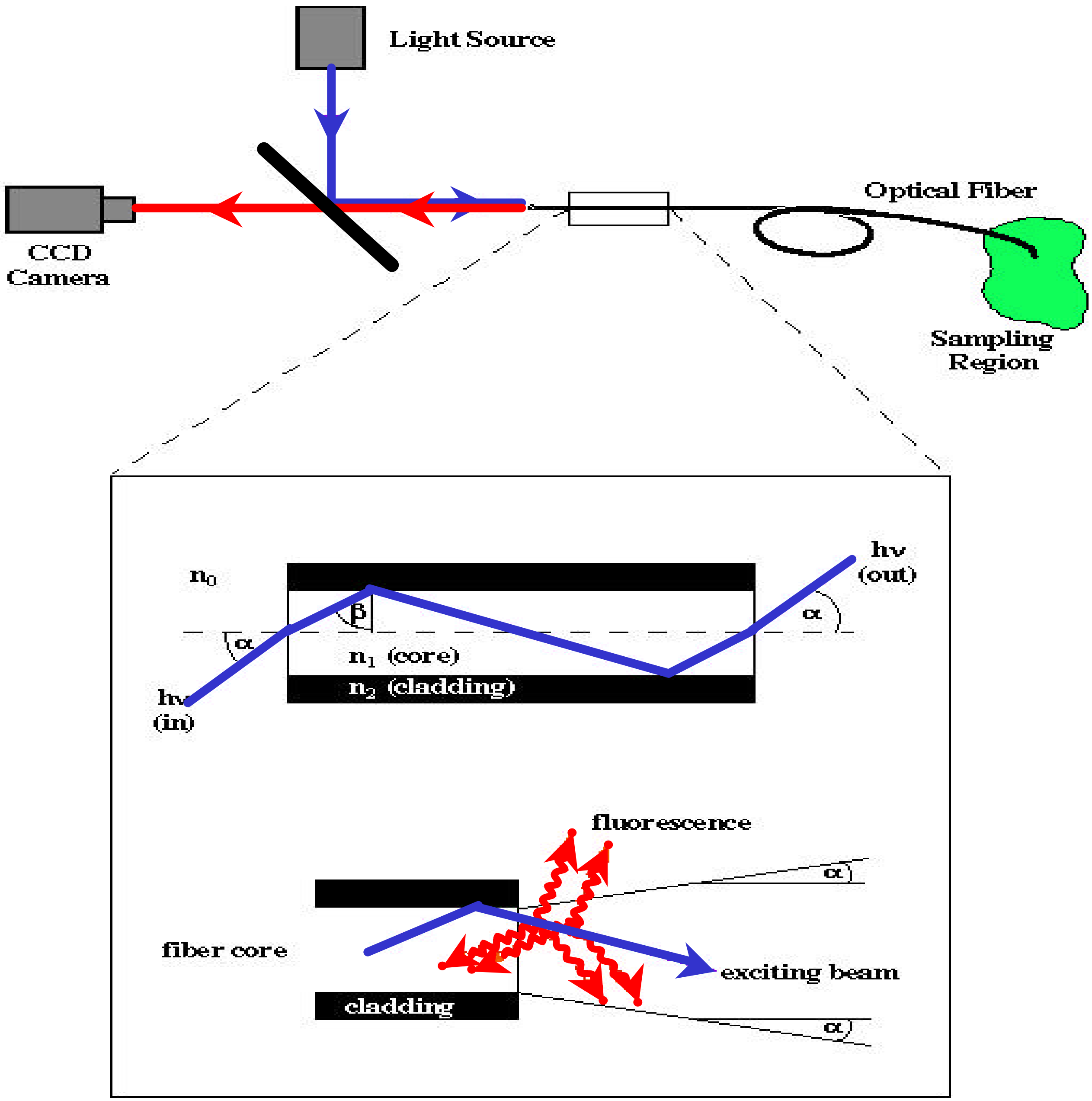
Tufts University

4 April 2002

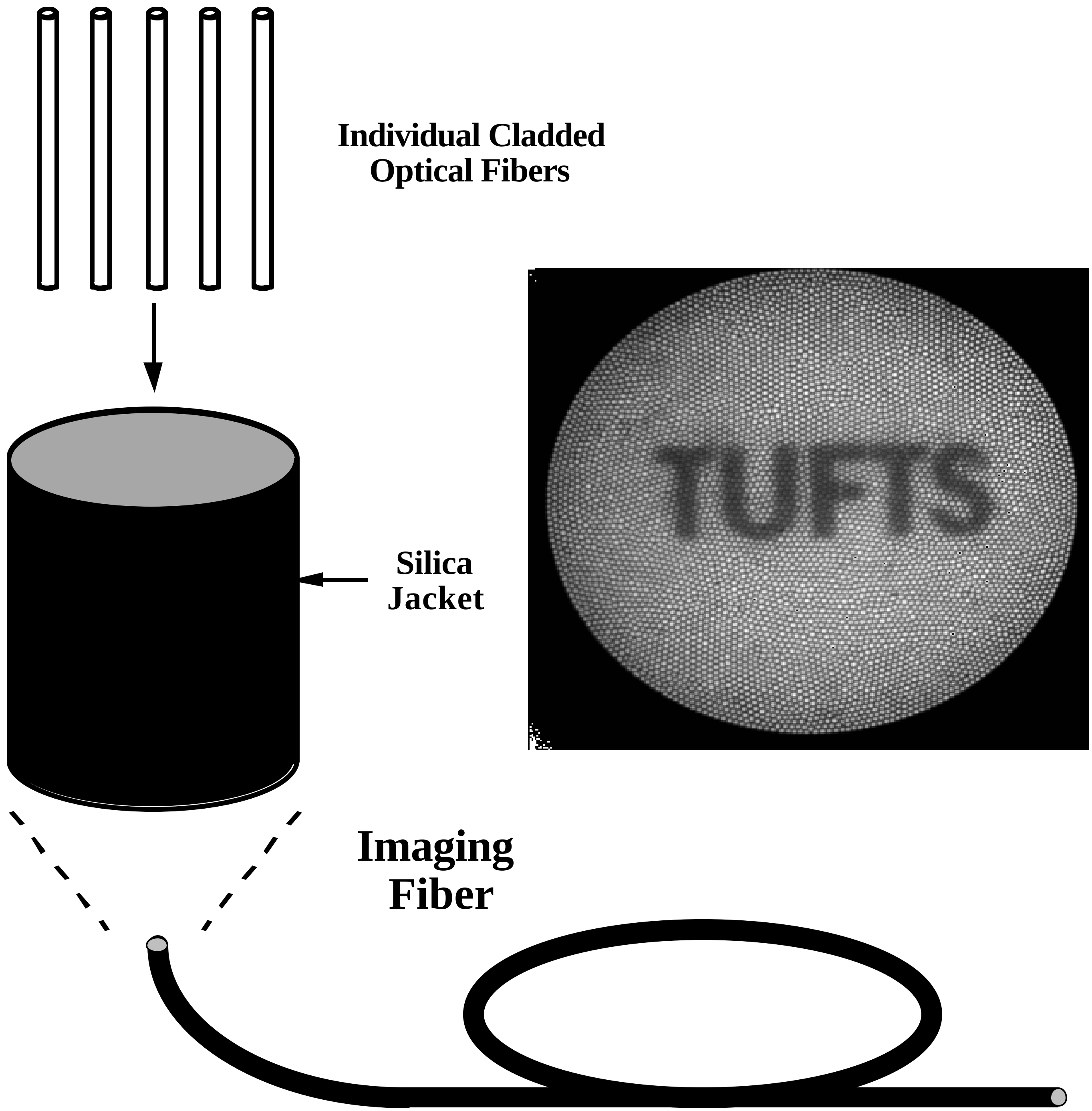
Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
04APR2002	N/A	03APR2002 - 04APR2002

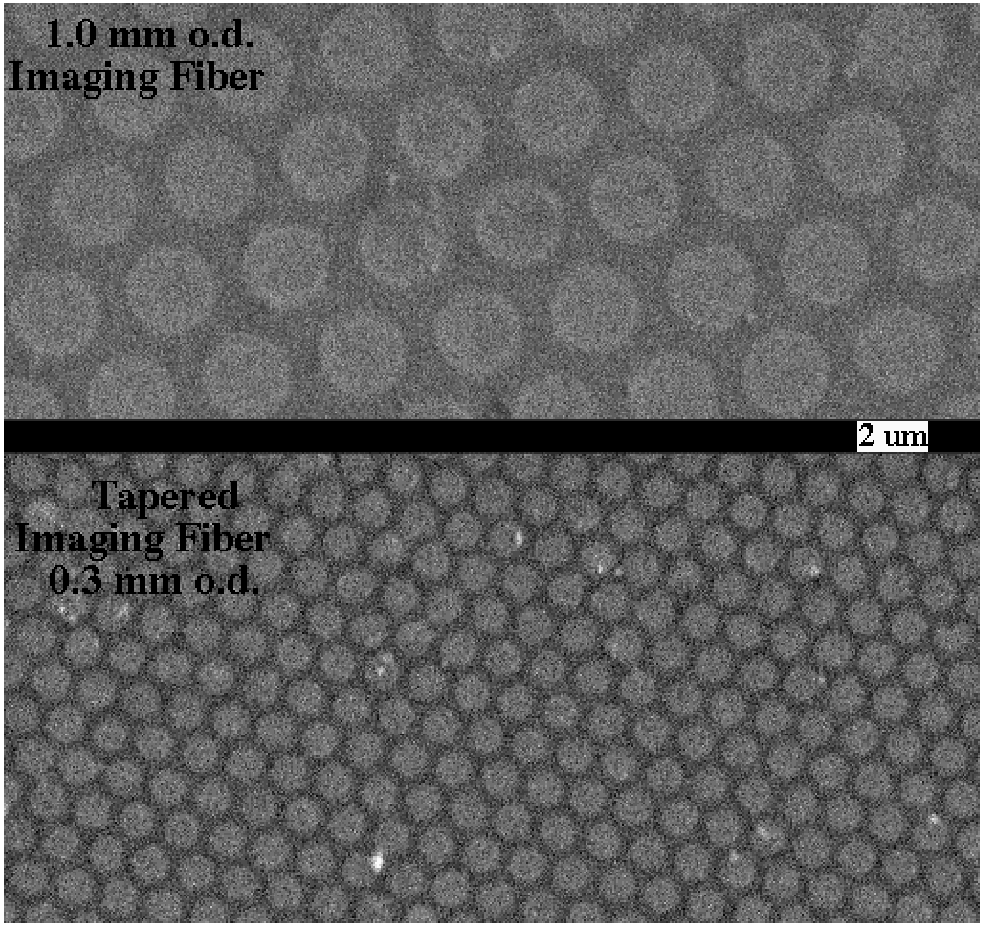
Title and Subtitle Microarrays	**Contract Number** F19628-00-C-0002
	Grant Number
	Program Element Number
Author(s) Walt, David	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Tufts University	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) Air Force ESC/XPK (Richard Axtell) Hanscom AFB, MA 01731	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Workshop paper from the New England Bioterrorism Preparedness Workshop held 3-4 april 2002 at MIT Lincoln Laboratory, Lexington, MA, The original document contains color images.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** SAR
Number of Pages 69	



Light Source
CCD
Camera
Optical Fiber
Sampling
Region
n_0
β
α
hv
(in)
hv
(out)
n_1 (core)
n_2 (cladding)
fluorescence
fiber core
exciting beam
cladding


Individual Cladded Optical Fibers
Silica Jacket
TUFTS
Imaging Fiber

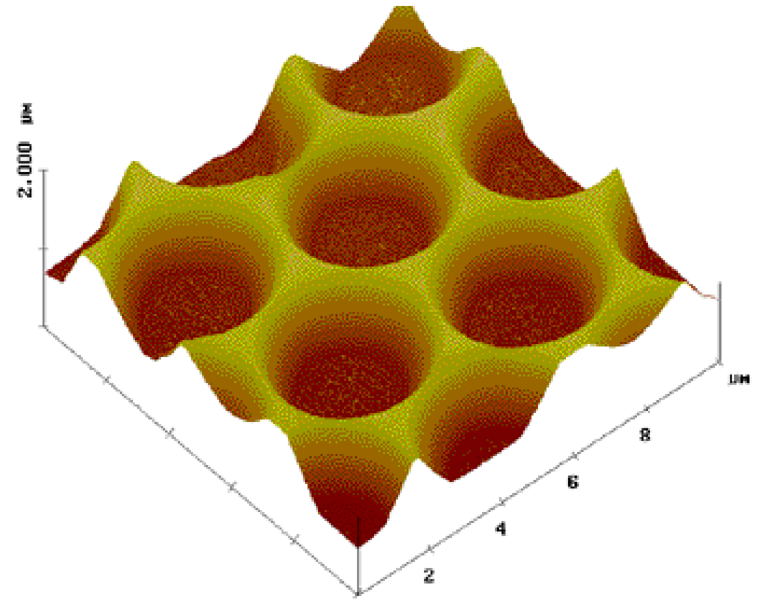
Optical Imaging Fiber Before and After Tapering



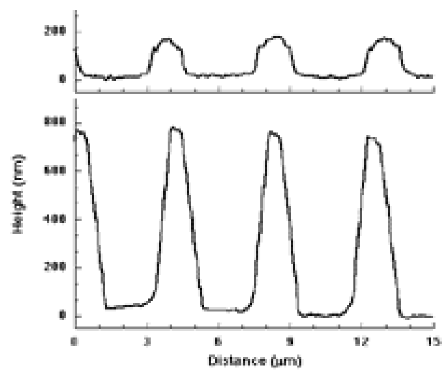
Individual Core Diameter ~ 2.6 ? m

Individual Core Diameter ~ 0.85 ? m

AFM of a Chemically-Etched 1000-?m Diameter Imaging Fiber



Well Profiles

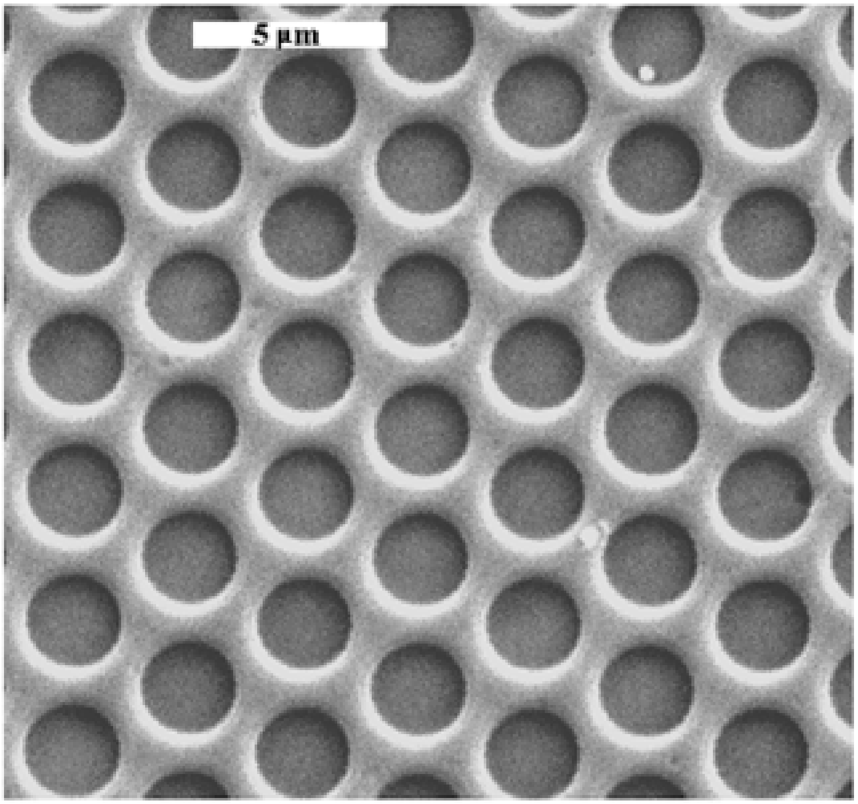


15 s etch

60 s etch

Pantano, P.; Walt, D.R. *Chem. Mater.* **1996**, 8, 2832-2835

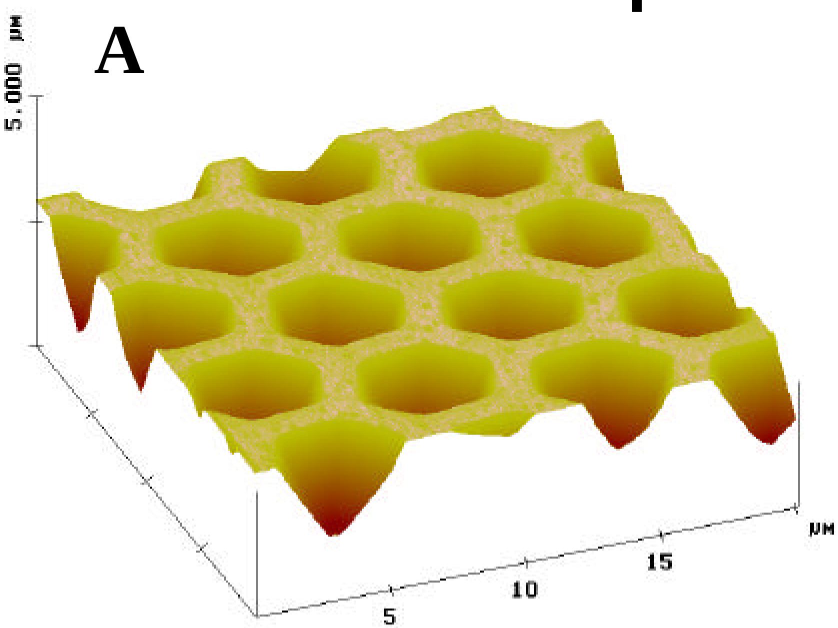
SEM of a Chemically-Etched 1000-? m Diameter Imaging Fiber

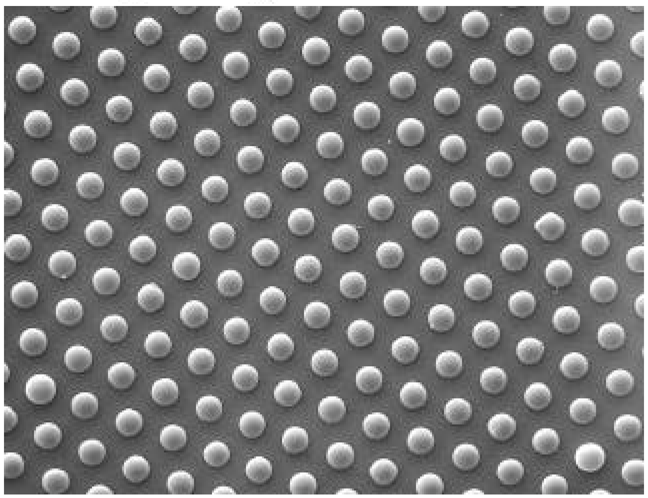


Pantano, P.; Walt, D.R. *Chem. Mater*. **1996**, 8, 2832-2835

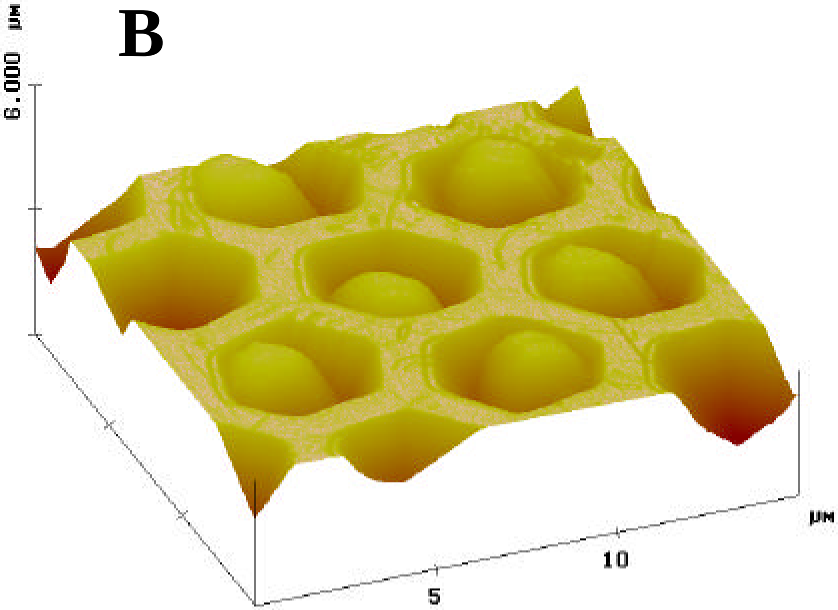
Microspheres in Microwells

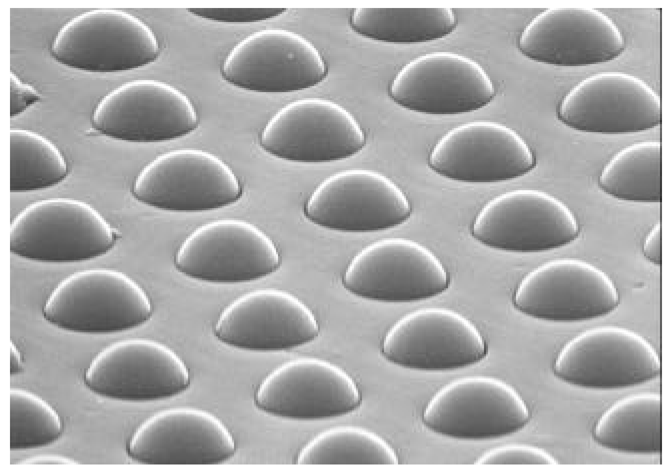
A



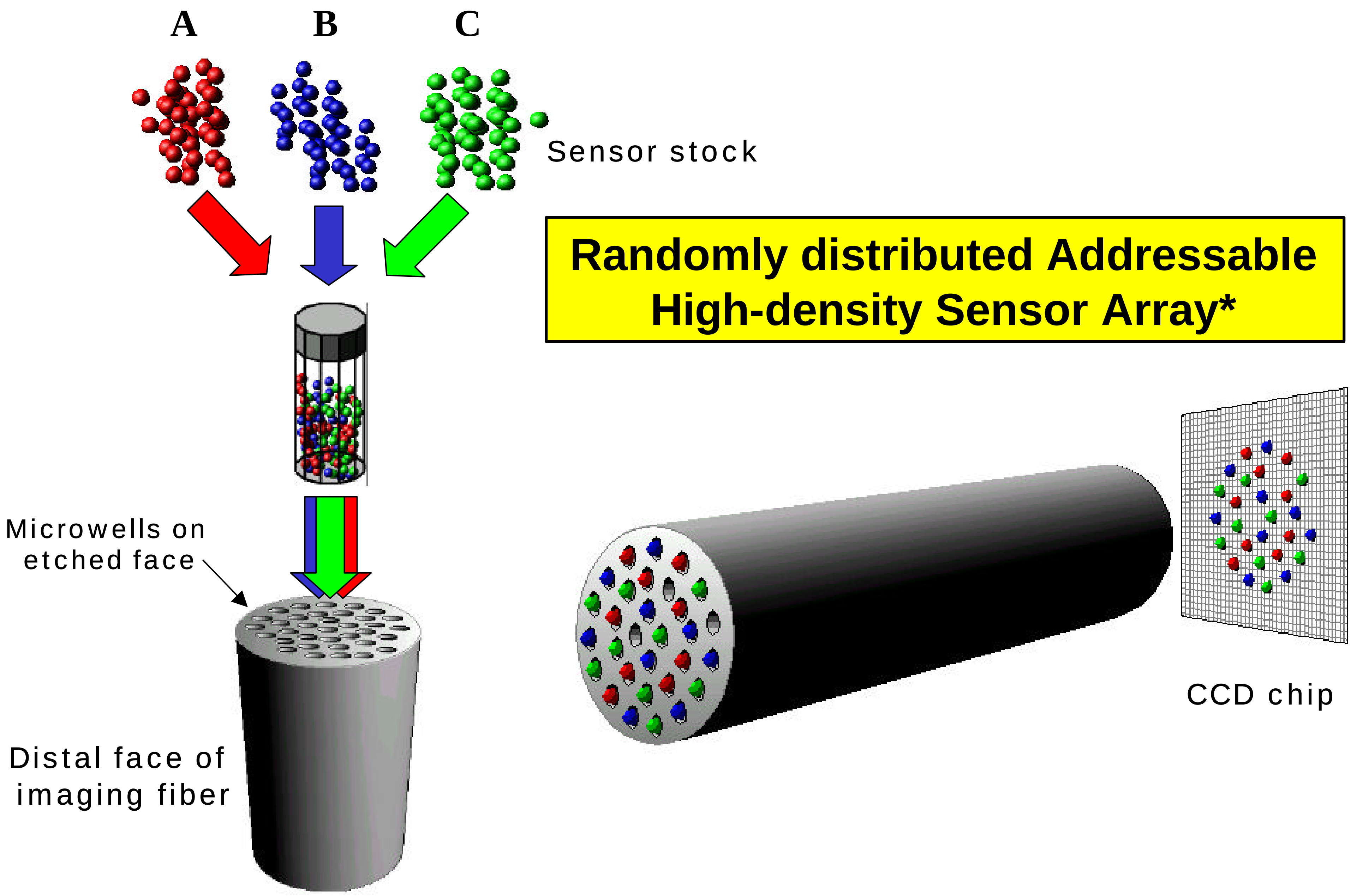


B





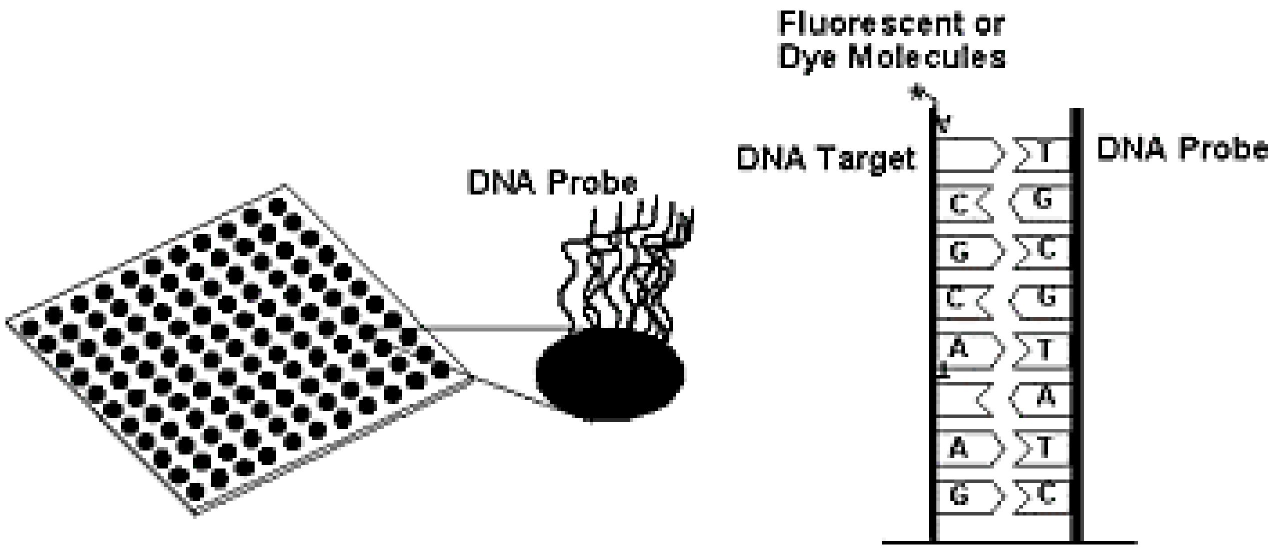
Michael, K.L. *et al. Anal. Chem.* 70 (7): 1242-1248 (1998).



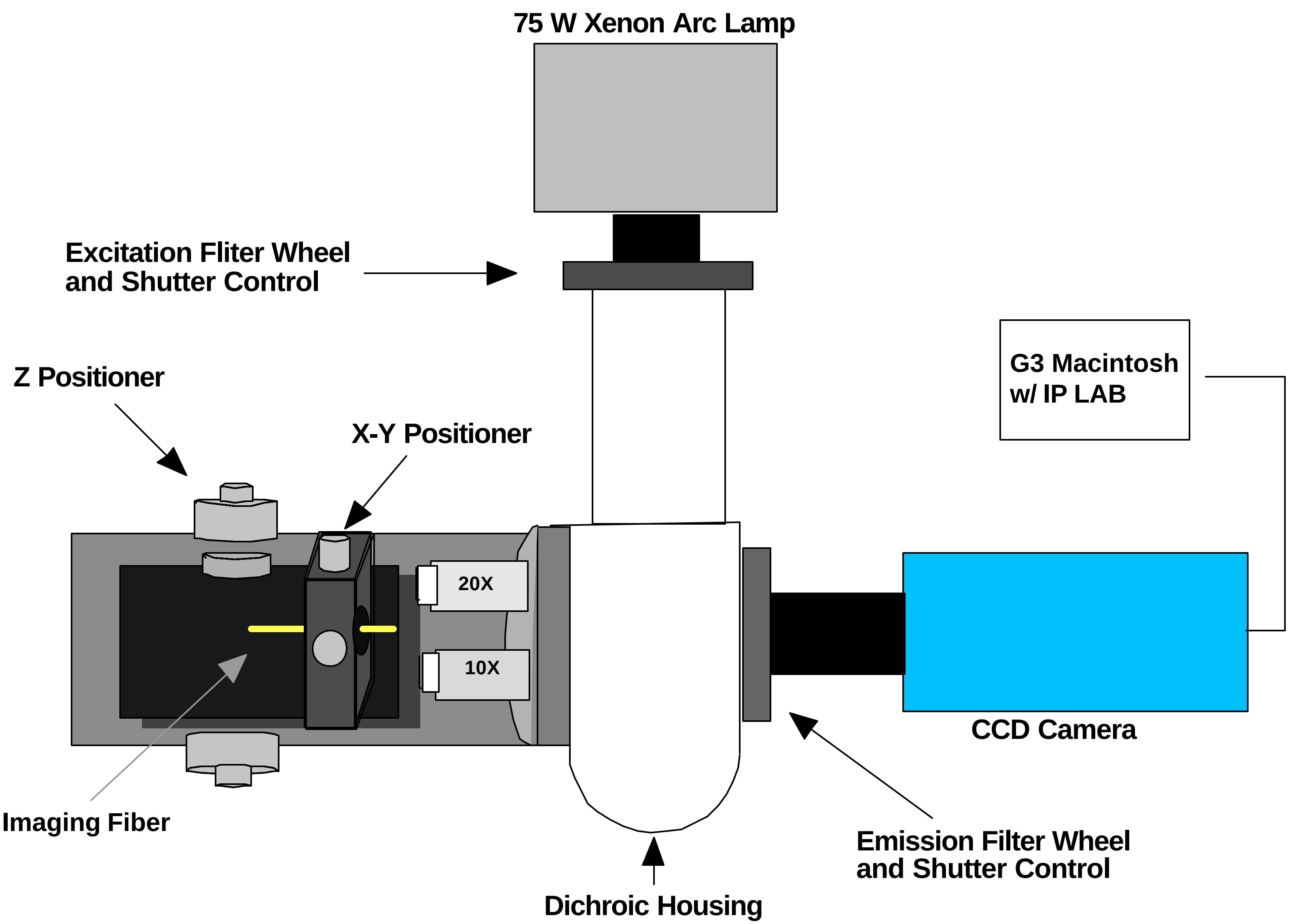
Randomly distributed Addressable High-density Sensor Array*

*Michael et al. **1998** *Anal. Chem.* ***70***: 1242-1248

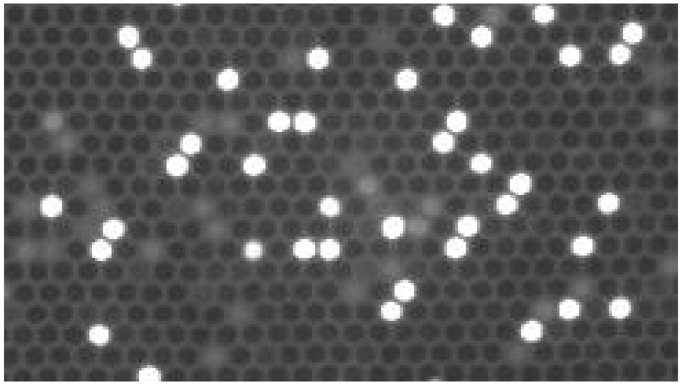
DNA Array Principle



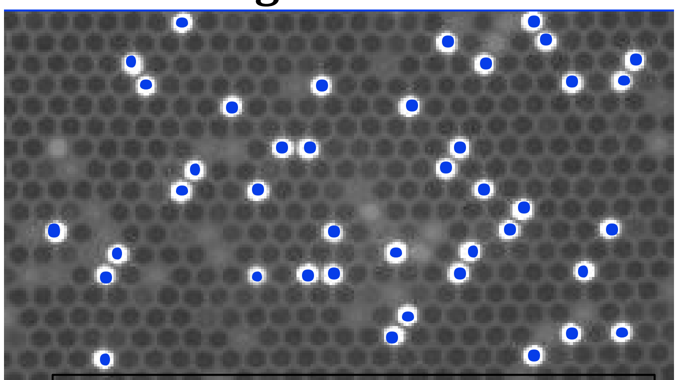
Instrumentation: Modified Fluorescence Microscope



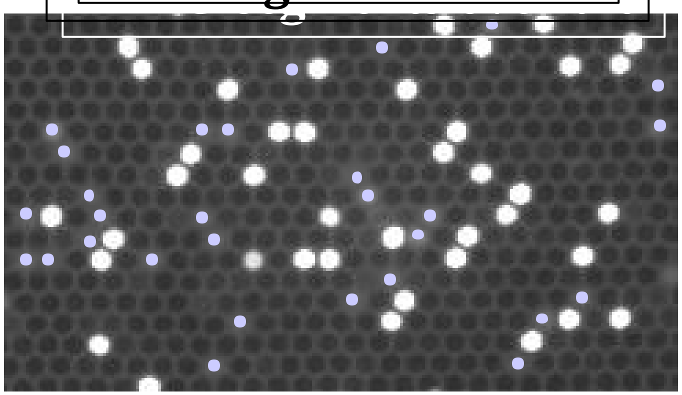
Encoding Signal of Dye 1



Signal 530

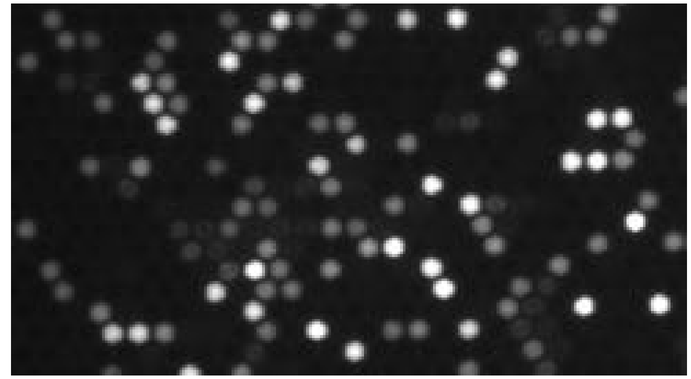


IFNG segments overlaid

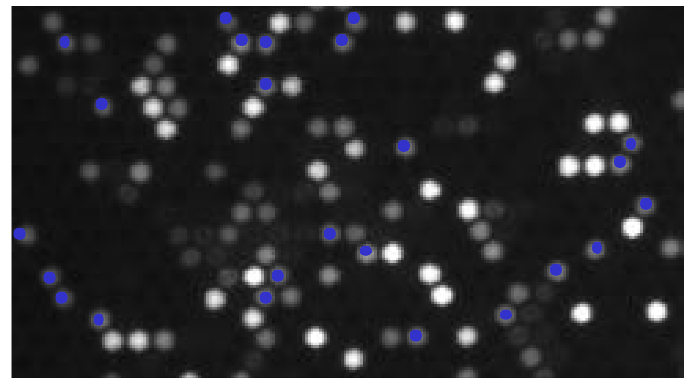


IL2 segments overlaid

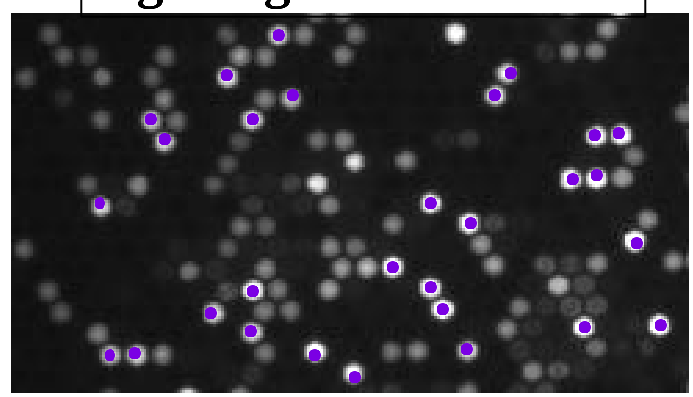
Encoding Signal of Dye 2



Signal 620



Bglo segments overlaid



Hwt segments overlaid

Sequences of 25 Probes used together in a Microsphere Array

1) ?-glo (segment of human ?-globin)[26]
TCA ACT TCA TCC ACG TTC ACC

2) IFNG (interferon gamma 1)[26]
IFNG TGG GTT CTC TTG GCT GTT ACT

3) IL2 (interleukin-2)[26]
TA CAA GAA TCC CAA ACT CAC CAG

4) IL4 (interleukin-4)[26]
CC AAC TGC TTC CCC CTC TGT

5) IL6 (interleukin-6)[26]
GT TGG GTC AGG GGT GGT TAT T

6) K-ras WT[27]
GGA GCT GGT GGC GTA

7) H-ras WT[27]
CCG GCG GTG T

8) CFTR (cystic fibrosis exon 11)[13]
CAT TAT ACT TGT AGA G

9) R553X (cystic fibrosis exon 10)[13]
TGT AGA ATT ATC TTC

10) PAN132[16] (human peripheral lymphocyte)
CCT CTA TAC TTT AAC GTC AAG

11) Schena-2[16]
AAG TTT AAC CTA TAC CCT GTC

12) Hakala-1[20]
CCT ATG ATG AAT ATA G

13) Hakala-2[20]
AAT ATG ATA ATG GCC T

14) complement to probe 1
TG AAC GTG GAT GAA GTT G

15) complement to probe 2
AG TAA CAG CCA AGA GAA CCC AAA

16) complement to probe 3
CT GGT GAG TTT GGG ATT CTT GTA

17) complement to probe 4
AC AGA GGG GGA AGC AGT TGG

18) complement to probe 5
AA TAA CCA CCC CTG ACC CAA C

19) complement to probe 6
TAC GCC ACC AGC TCC

20) complement to probe 7
ACA CCG CCG G

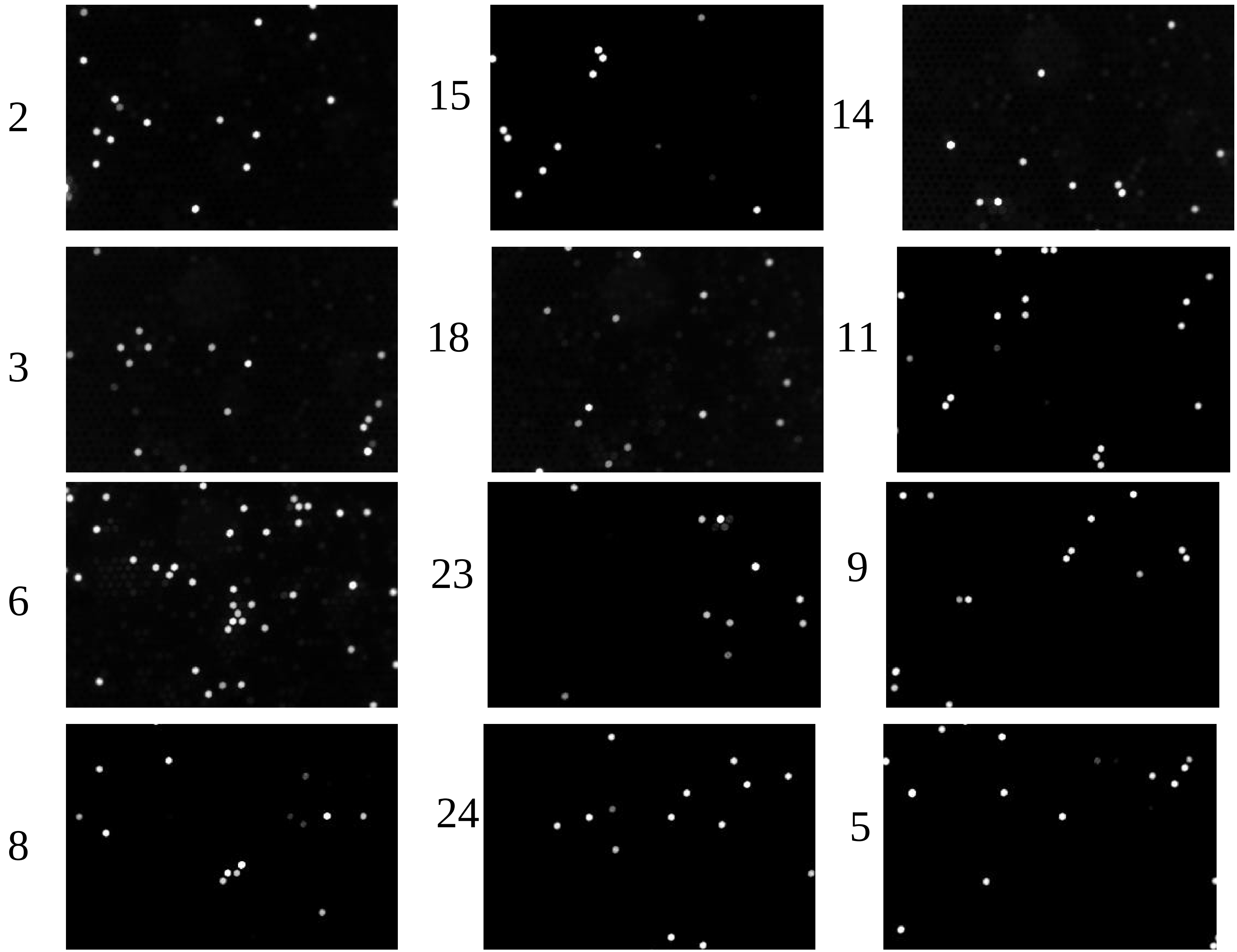
21) complement to probe 8
CTC TAC AAG TAT AAT G

22) complement to probe 9
GAA GAT GTT AAA GTA TAG AGG

23) complement to probe 10
CTA GAC GTT AAA GTA TAG AGG

24) complement to probe 12
CTA TAT TCA TCA TAG G

25) complement to probe 13
AGG CCA TTA TCA TAT T


2
15
14
3
18
11
6
23
9
8
24
5

E. coli Allelic Discrimination

ycgW locus*

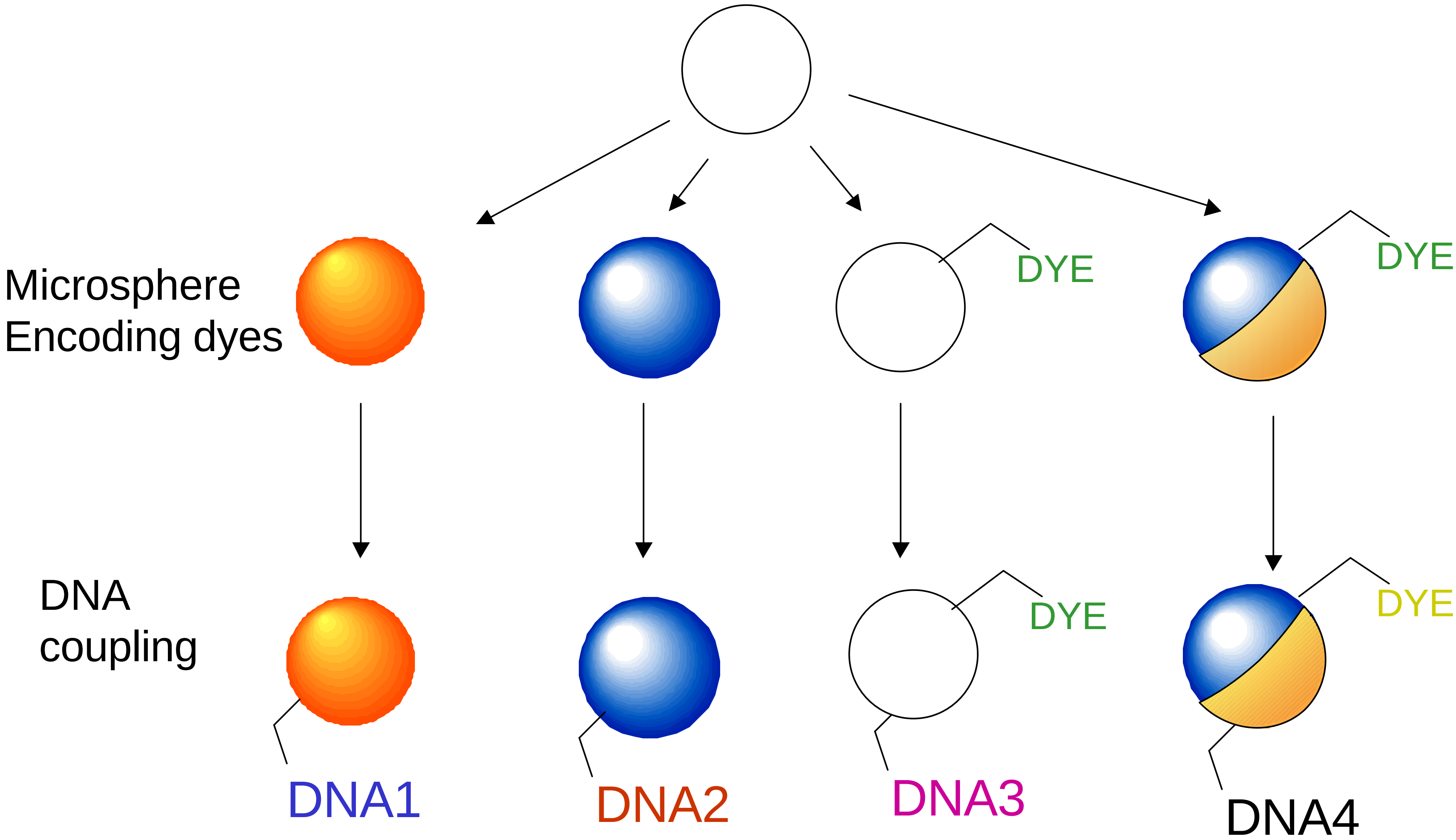
No.	strain	strain sequence											
1	BSR9b	T	T	T	T	T	G	A	A	G	G	G	G
2	BSR9c	T	T	T	T	T	G	A	A	G	G	G	G
3	"ETEC"	T	T	T	T	T	T	G	A	G	G	G	G
4	O111NM	T	T	T	T	T	G	A	A	G	G	G	G
5	O113:H2	T	T	T	T	T	G	A	A	G	G	G	G
6	O157NM	T	T	T	T	T	T	G	A	G	G	G	G
7	HER1058	no PCR product											
8	K12DH5a	T	T	T	T	T	T	G	A	G	G	G	G
9	K12W4100	T	T	T	T	T	T	G	A	G	G	G	G
10	O55:H7	T	T	T	T	C	G	A	A	G	G	G	G
11	"EPEC"	T	T	T	T	T	T	G	A	G	G	G	G
12	K12W3110	T	T	T	T	T	T	G	A	G	G	G	G
13	O22:H8	T	T	T	T	T	G	A	A	G	G	G	G
14	O26:H-	T	T	T	T	T	T	G	A	G	G	G	G
15	O42:H2	T	T	T	T	T	G	A	A	G	G	G	G
16	O157:H7	no PCR product											
17	HER1057	no PCR product											
18	HER1261	no PCR product											
	Cons.	T	T	T	*	*	*	*	*	*	*	*	G

No.	strain	strain sequence											
19	HER1265	no PCR product											
20	HER1266	no PCR product											
21	EC68	no PCR product											
22	EC69	T	T	T	T	T	T	G	A	G	G	G	G
23	EC63	no PCR product											
24	EC54	no PCR product											
27	O86:H10	T	T	T	T	T	G	A	A	G	G	G	G
37	O86:H18	T	T	T	G	T	T	T	T	T	T	T	G
30	O8:H9	T	T	T	T	T	G	A	A	G	G	G	G
34	O9:H33	T	T	T	T	T	T	G	A	G	G	A	G
38	O153:H-	T	T	T	T	T	T	G	A	G	G	G	G
43	O26:H11	T	T	T	T	T	G	A	A	G	G	G	G
48	O127:H21	T	T	T	T	T	G	A	A	G	G	G	G
52	EC1	T	T	T	T	T	T	G	A	G	G	G	G
53	EC7	T	T	T	T	T	T	G	A	G	G	G	G
54	EC18	T	T	T	T	T	T	G	A	G	G	A	G
55	EC47	T	T	T	G	T	T	T	T	T	T	G	G
56	EC52	no PCR product											
	Cons.	T	T	T	*	*	*	*	*	*	*	*	G

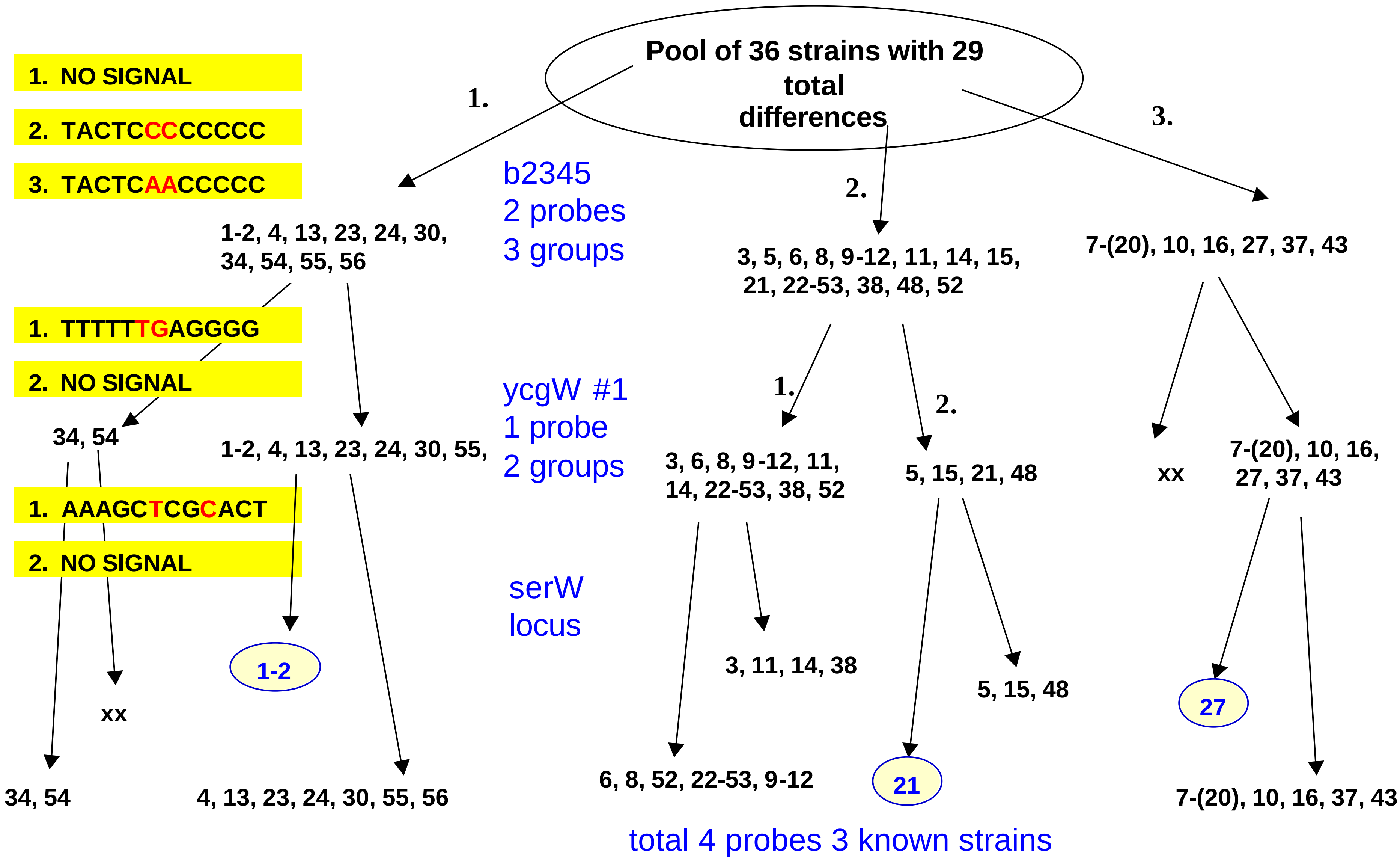
= probe 1 signal   = no signal

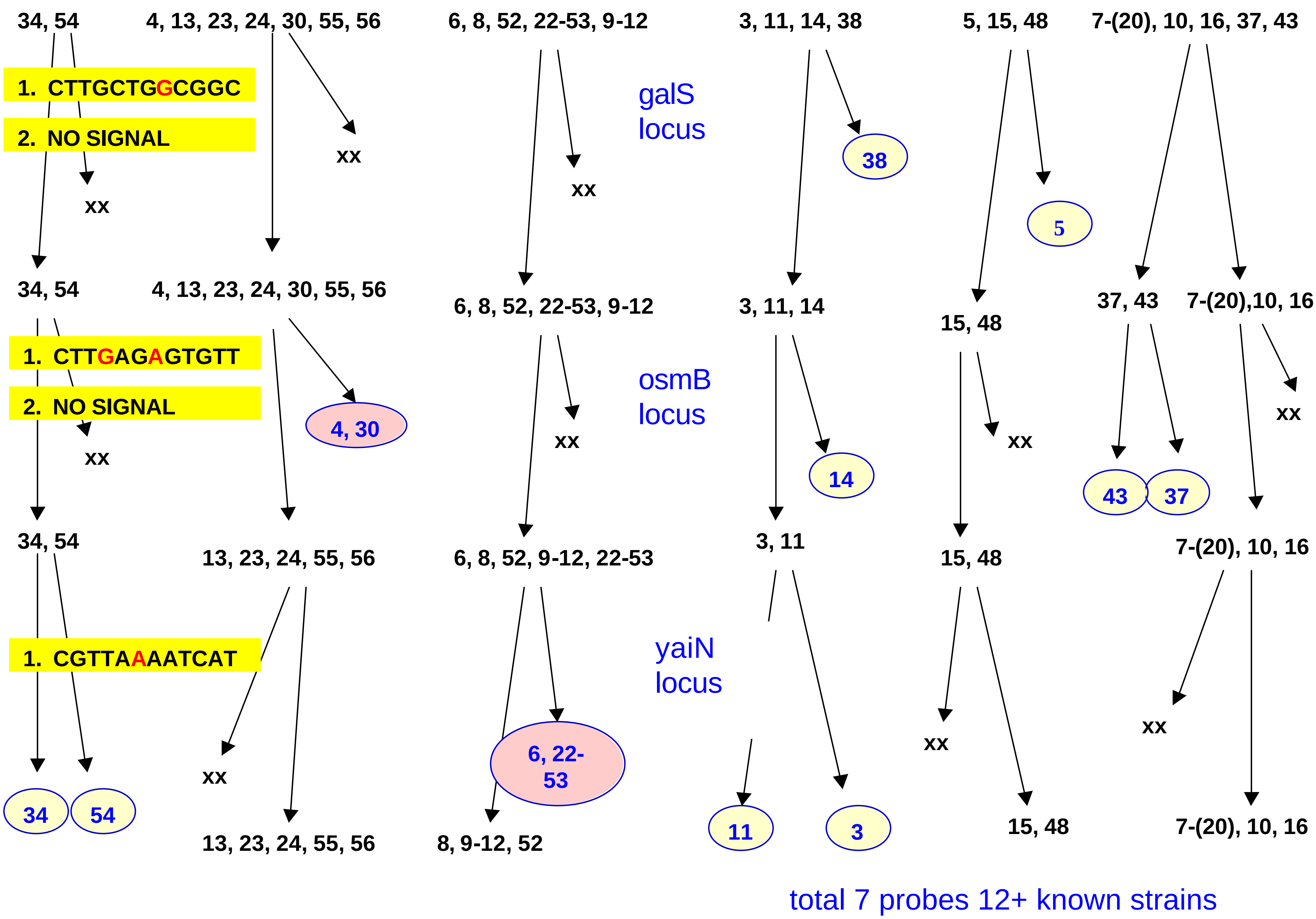
*ycgW locus is 77 nucleotides long

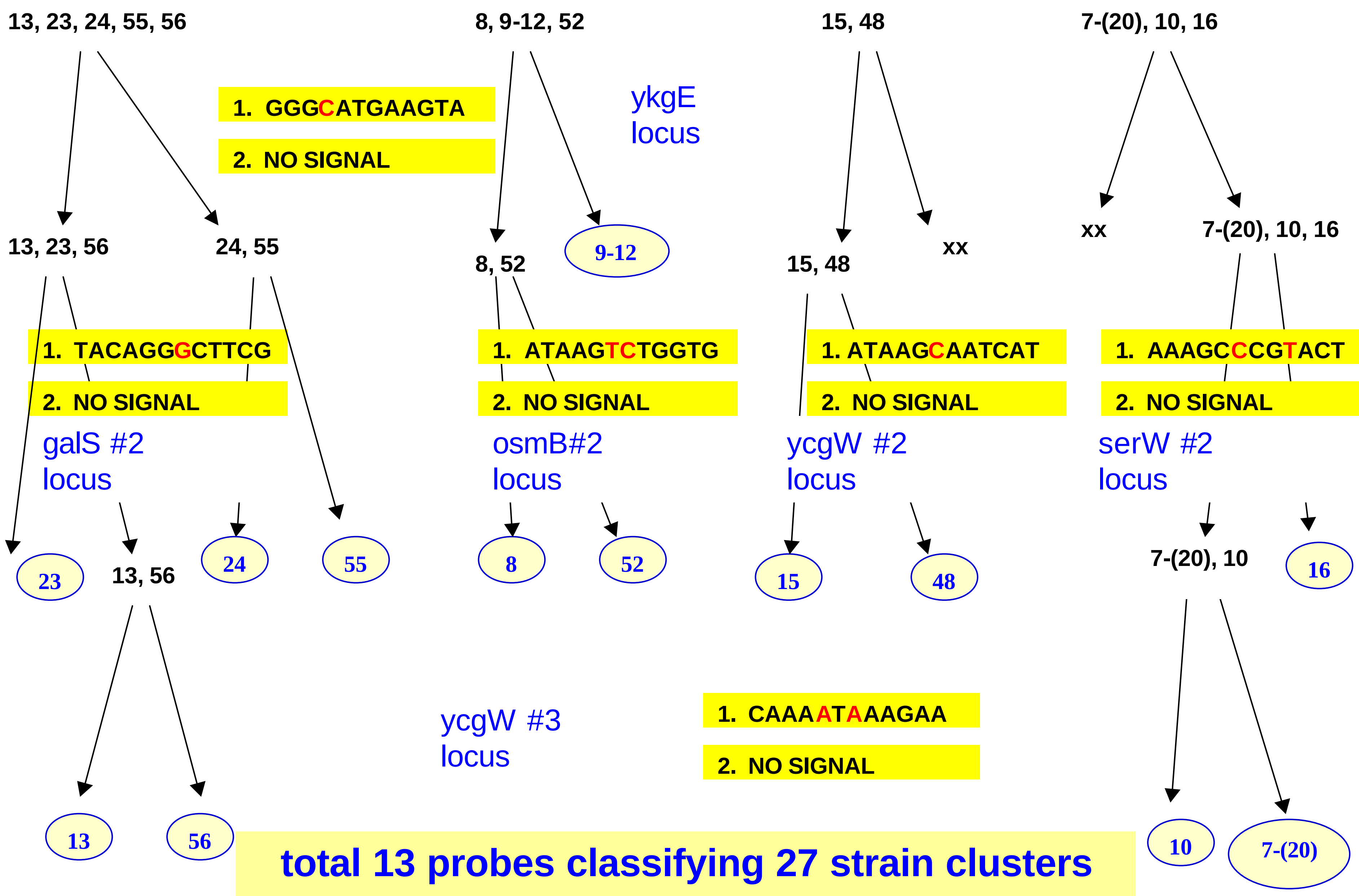
Microsphere Functionalization



E. coli Genomic Discrimination Flowchart




34, 54
4, 13, 23, 24, 30, 55, 56
6, 8, 52, 22-53, 9-12
3, 11, 14, 38
5, 15, 48
7-(20), 10, 16, 37, 43
1. CTTGCTGGCGGC
2. NO SIGNAL
xx
galS
locus
38
5
34, 54
4, 13, 23, 24, 30, 55, 56
6, 8, 52, 22-53, 9-12
3, 11, 14
15, 48
37, 43
7-(20),10, 16
1. CTTGAGAGTGTT
2. NO SIGNAL
4, 30
osmB
locus
14
43
37
34, 54
13, 23, 24, 55, 56
6, 8, 52, 9-12, 22-53
3, 11
15, 48
7-(20), 10, 16
1. CGTTAAAATCAT
yaiN
locus
6, 22-53
34
54
13, 23, 24, 55, 56
8, 9-12, 52
11
3
15, 48
7-(20), 10, 16
total 7 probes 12+ known strains



13, 23, 24, 55, 56
1. GGGCATGAAGTA
2. NO SIGNAL
13, 23, 56
24, 55
1. TACAGGGCTTCG
2. NO SIGNAL
galS #2 locus
23
13, 56
24
55
13
56
8, 9-12, 52
ykgE locus
8, 52
9-12
1. ATAAGTCTGGTG
2. NO SIGNAL
osmB#2 locus
8
52
15, 48
15, 48
xx
1. ATAAGCAATCAT
2. NO SIGNAL
ycgW #2 locus
15
48
ycgW #3 locus
1. CAAAATAAAGAA
2. NO SIGNAL
7-(20), 10, 16
xx
7-(20), 10, 16
1. AAAGCCCGTACT
2. NO SIGNAL
serW #2 locus
7-(20), 10
16
10
7-(20)
total 13 probes classifying 27 strain clusters

E. coli Genomic Pattern Response

No.	strain	ycgW	serW	osmB	yaiN	ykgE
14	O26:H-					
11	"EPEC"					
3	"ETEC"					
6	O157NM					
22	EC69					
9	K12W4100					
12	K12W3110					
8	K12DH5a					

= signal response

= no signal

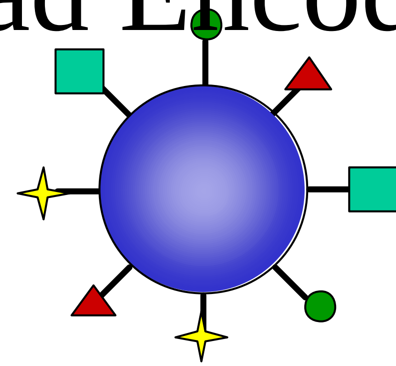
E. coli Genomic Pattern Response

No.	strain	ycgW	b2345	serW	galS	Osmb	yKgE	ycgW#3	serW#2
21	EC68								
1	BSR9b								
2	BSR9c								
15	O42:H2								
5	O113:H2								
4	O111NM								
24	EC54								
13	O22:H8								
23	EC63								
10	O55:H7								
16	O157:H7								
7	HER1058								
17	HER1057								
18	HER1261								
19	HER1265								
20	HER1266								

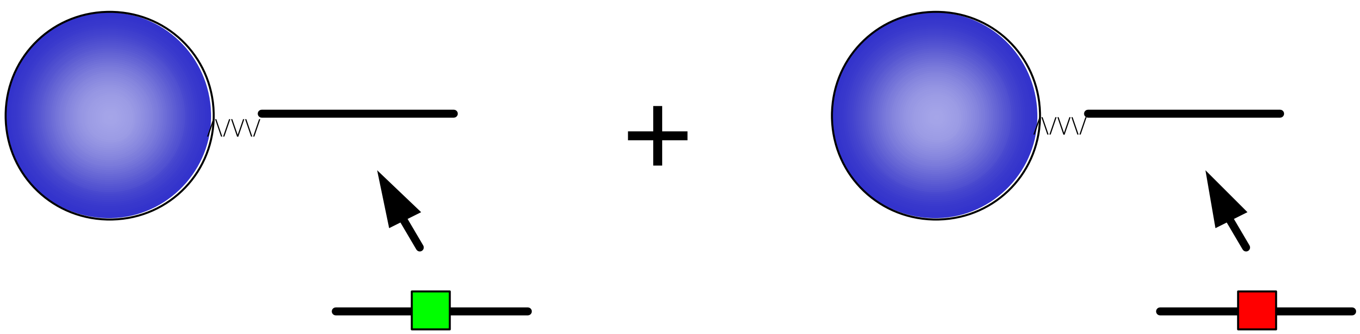
= signal response

= no signal

Bead Encoding

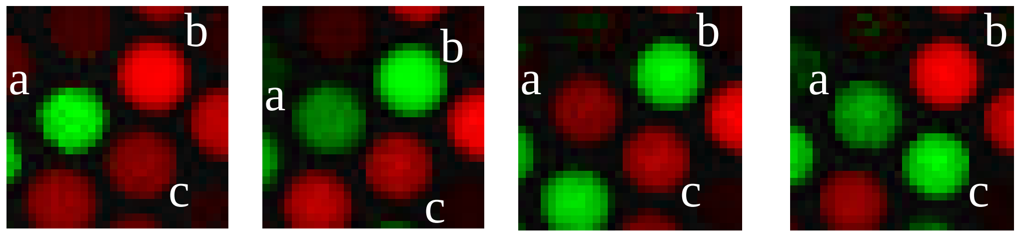


Sequential Decoding

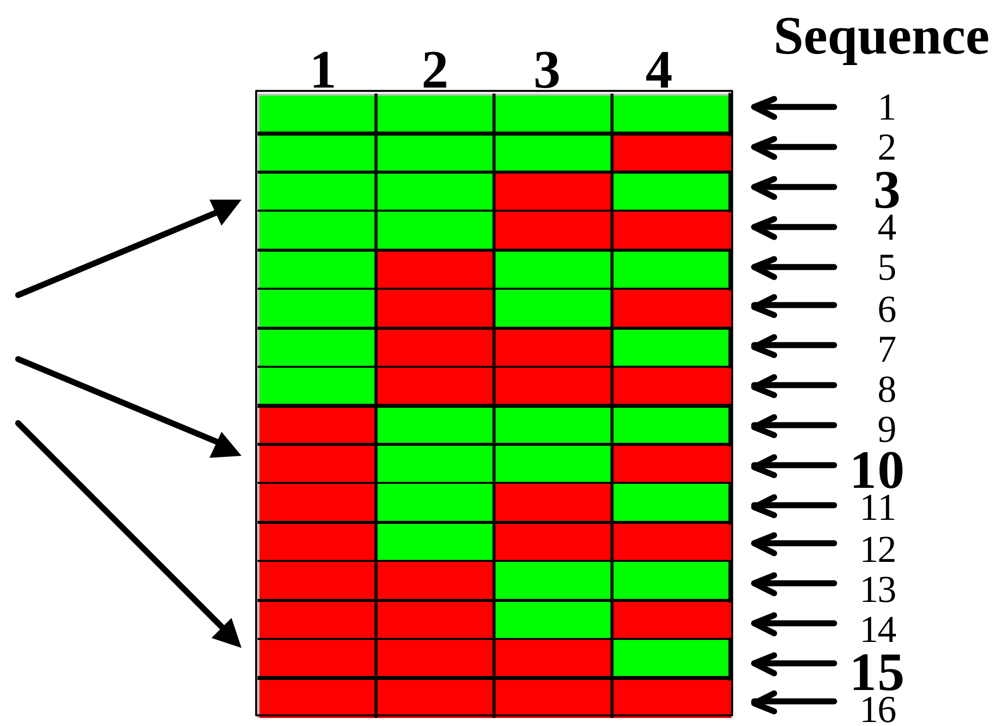


Decoding 16 Probes

Hyb # 1 2 3 4



	Hyb.#1	Hyb.#2	Hyb.#3	Hyb.#4
a	**green**	**green**	**red**	**green**
b	**red**	**green**	**green**	**red**
c	**red**	**red**	**red**	**green**

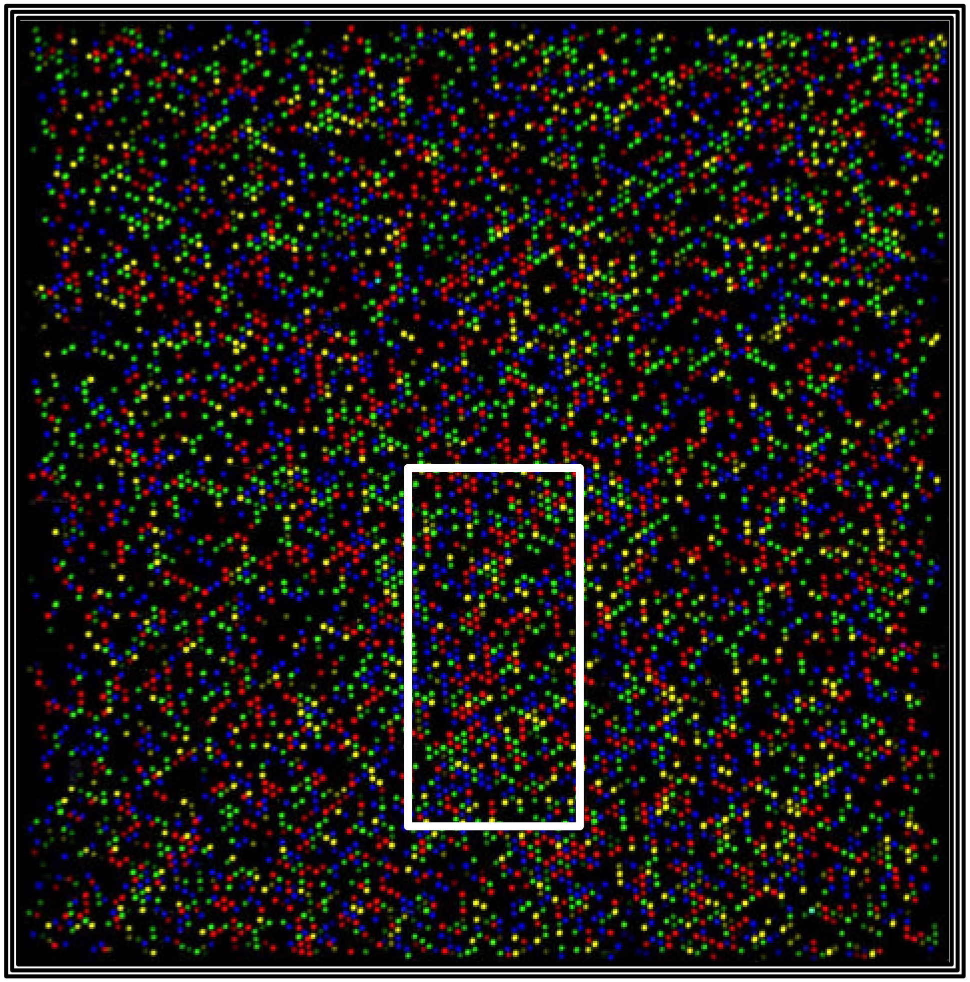


Decoding is Exponential

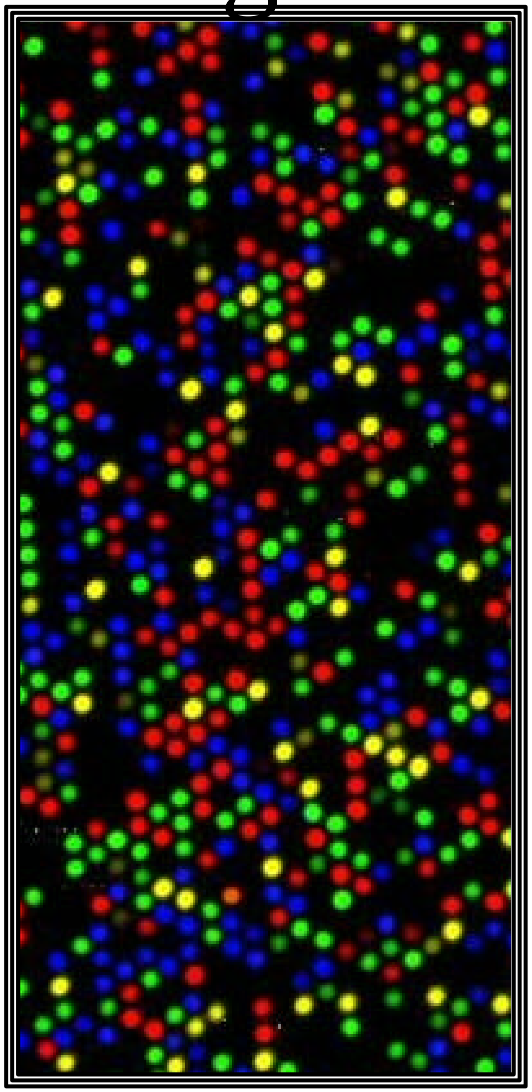
***x* labels , *z* steps = x^z codes**

- **2** Dyes ^ **4** Steps = **16** Codes
- **4** Dyes ^ **6** Steps = **4,096** Codes

Four-Color Decoding

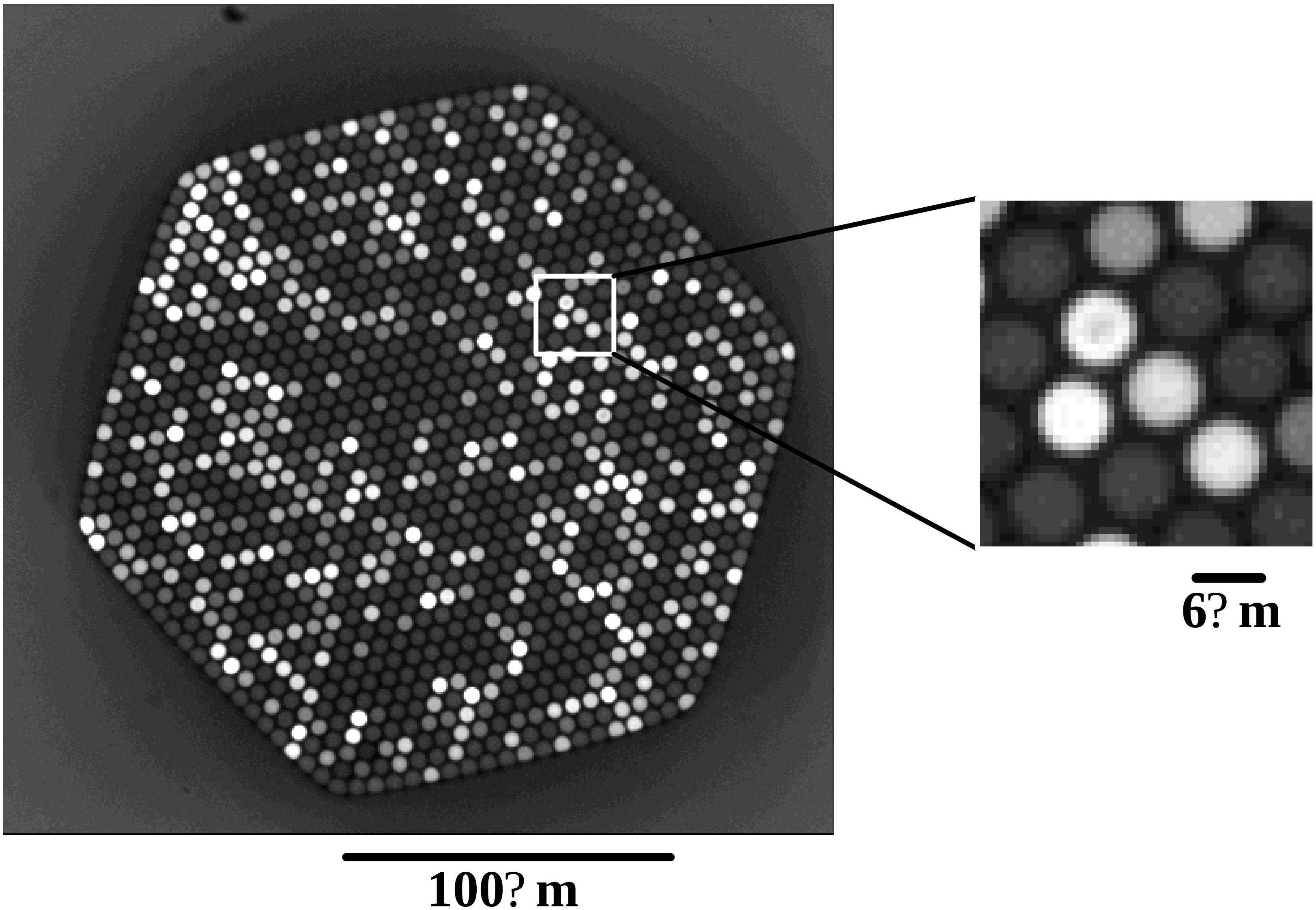


~13,000 Wells, 16 Probe Sequences

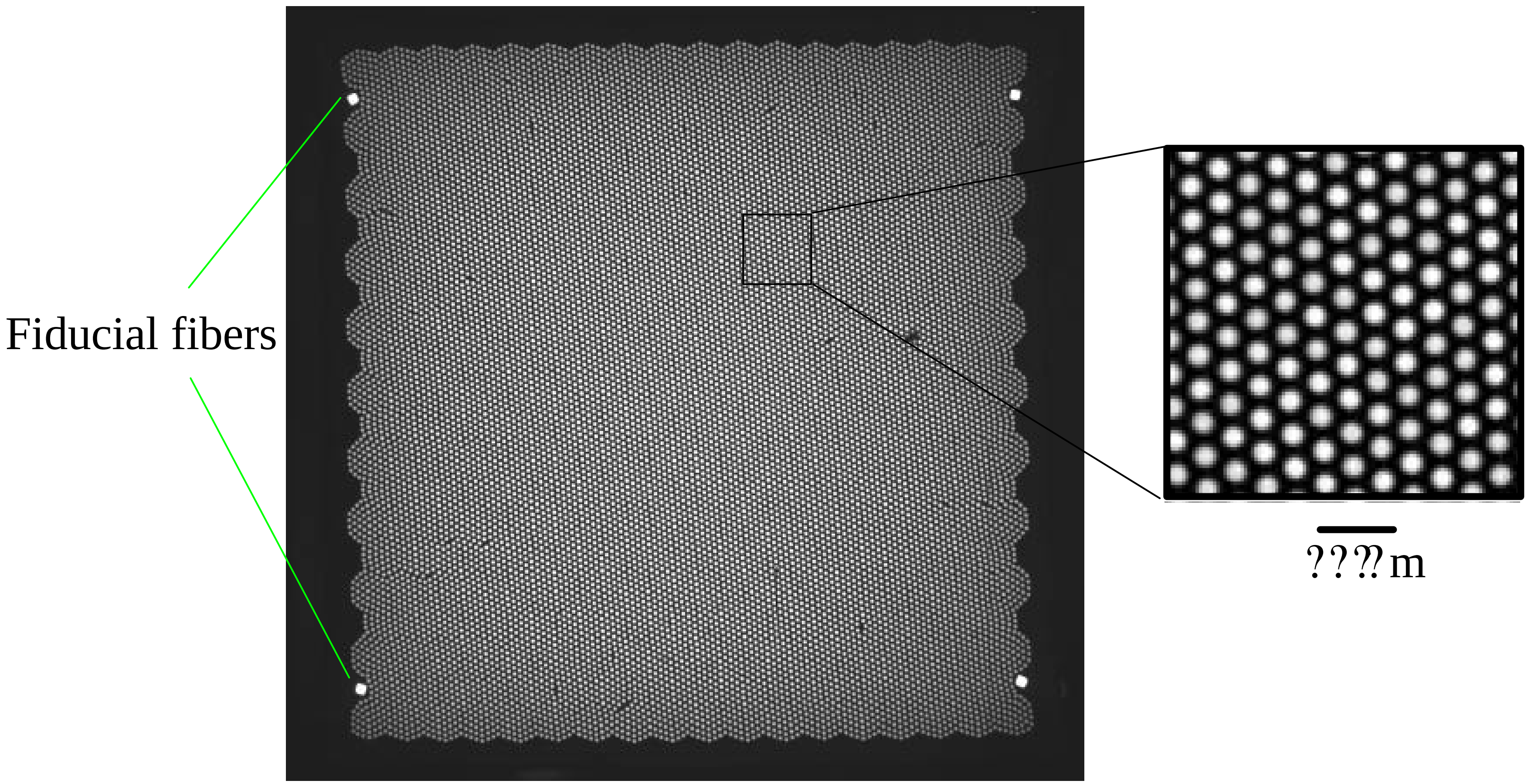


D.R. Walt, *Science*, 2000

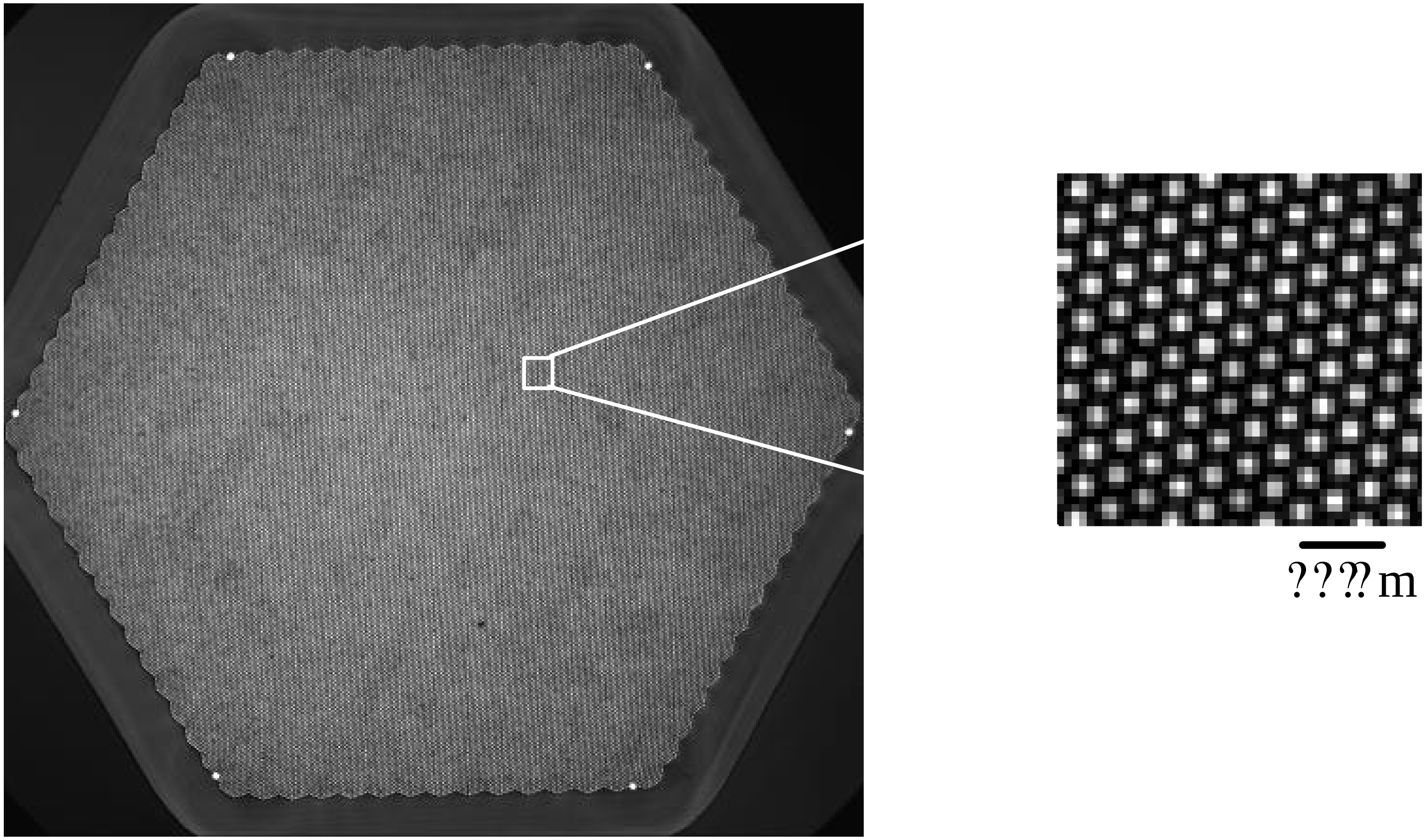
1K Fiber Bundle



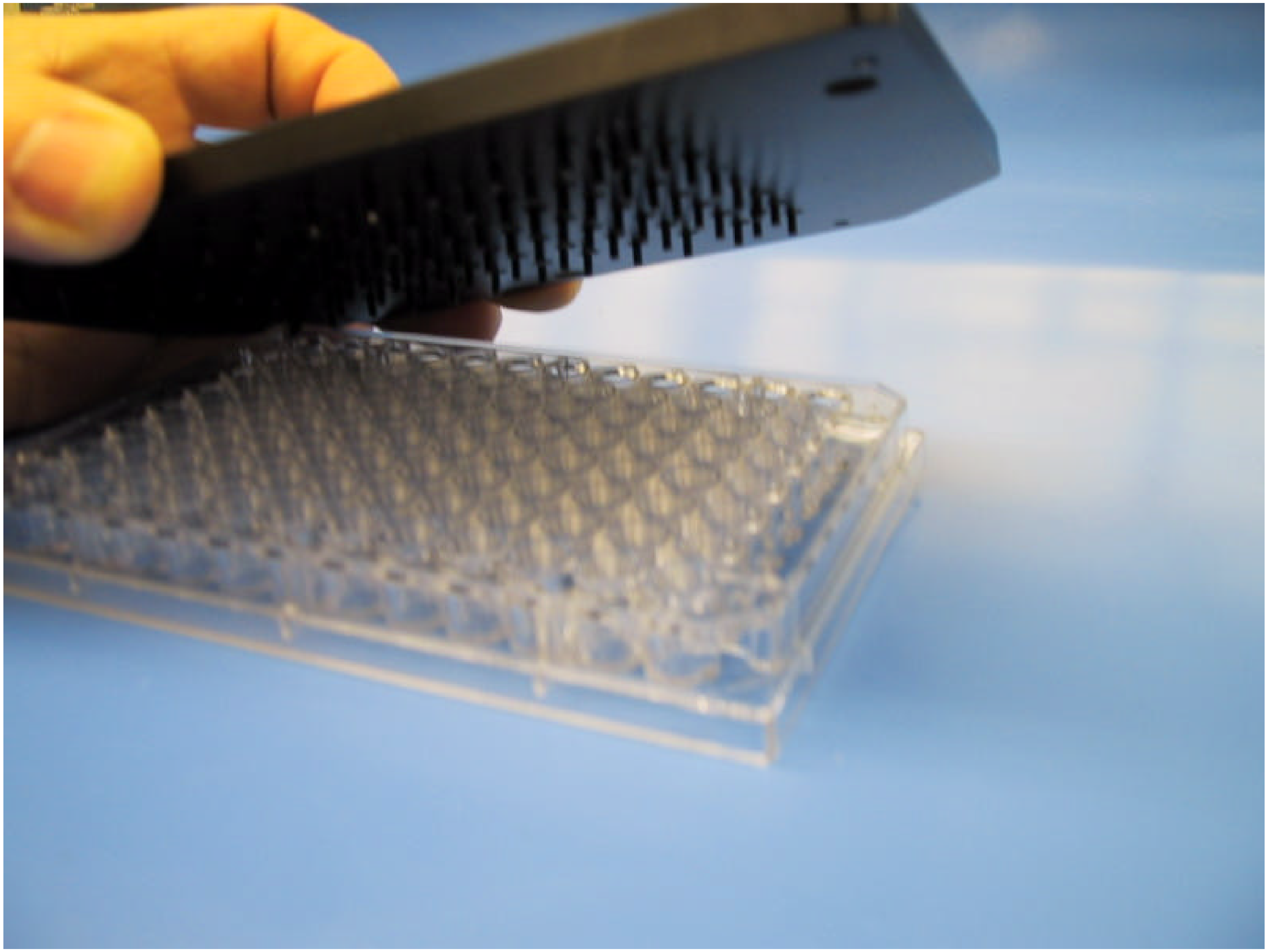
13K Fiber Bundle



57K Fiber Bundle



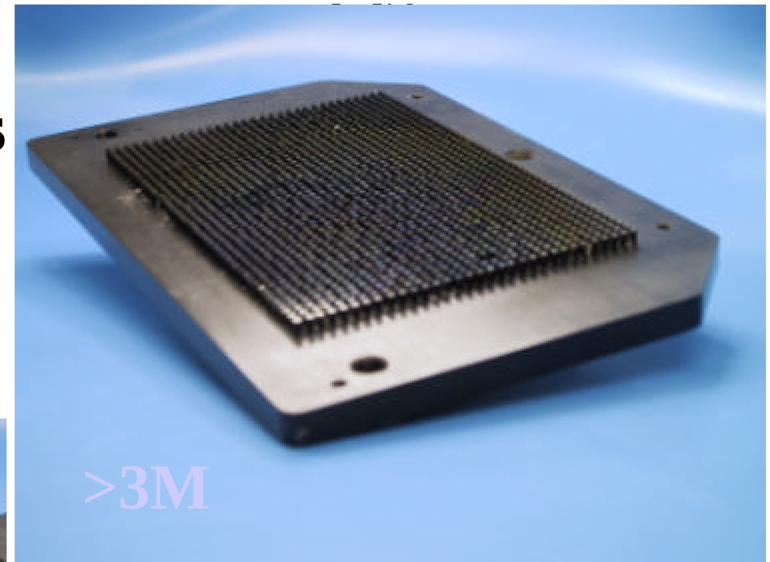
Array of Arrays™

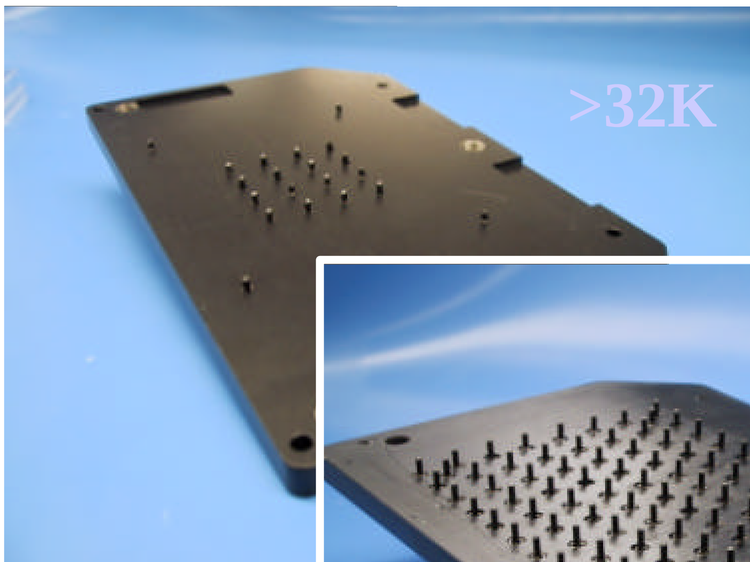


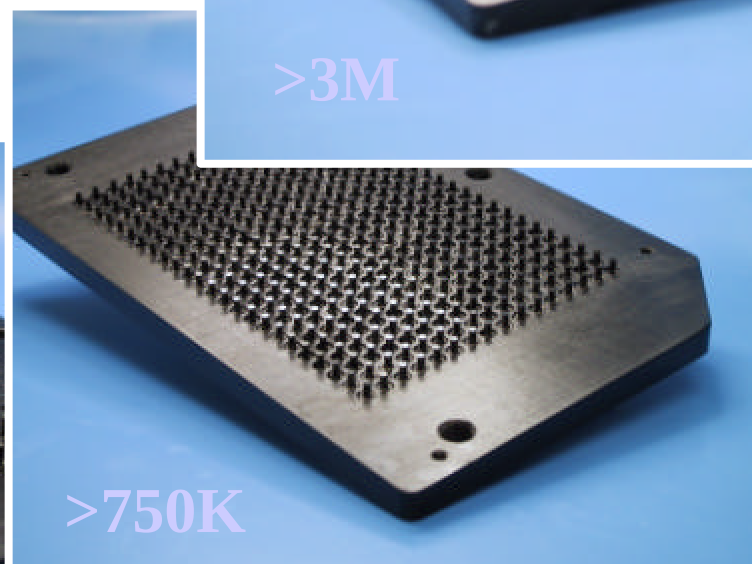
Scalability of Technology

Unique Experiments

(with ~20-fold redundancy)







Bead Summing

Concentration	Hybridization Time: No Summing	Hybridization Time: 100 Bead Summed
100 pM	10 minutes	4 minutes
10 pM	30 minutes	7 minutes
100 fM	4 hours	20 minutes
10 fM	17 hours	30 minutes

Size and Concentration

Volume			1?? M	1 nM	1 pM
(1 mm)^3	???1?? L	10^{-6} L	$6x10^{11}$	$6x10^{8}$	$6x10^{5}$
(100 ? m)^3	1 nL	10^{-9} L	$6x10^{8}$	$6x10^{5}$	$6x10^{2}$
(10 ? m)^3	1 pL	10^{-12} L	$6x10^{5}$	$6x10^{2}$	$6x10^{-1}$
(1 ? m)^3	1 fL	10^{-15} L	$6x10^{2}$	$6x10^{-1}$	
(0.1 ? m)^3	1 aL	10^{-18} L	$6x10^{-1}$		

Probe and Target Sequences for DNA Microarray Detection Limits

Probe

IL2 (interleuken-2)	5’-TA-CAA-GAA-TCC-CAA-ACT-CAC-CAG-3'
IL6 (interleuken-6)	5’-GT-TGG-GTC-AGG-GGT-GGT-TAT-T-3’
F508C	5’-TAG-GAA-ACA-CCA-CAG-ATG-ATA-3’

Target

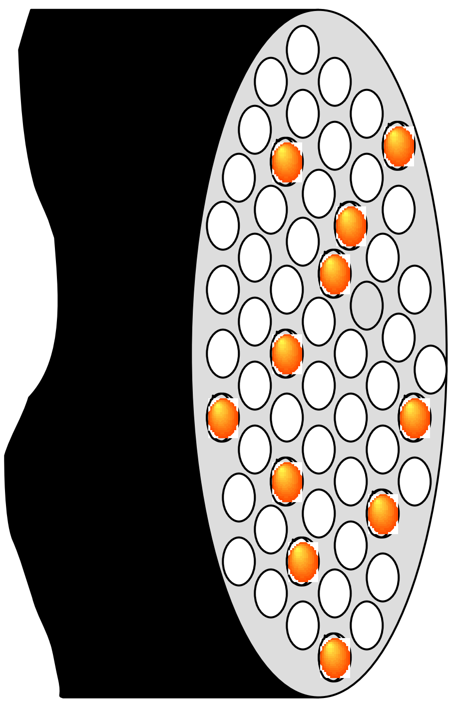
IL2 (interleuken-2)	5’-CT-GGT-GAG-TTT-GGG-ATT-CTT-GTA-3’
IL6 (interleuken-6)	5’-AA-TAA-CCA-CCC-CTG-ACC-CAA-C-3’
F508C	5’-TA-TCA-TCT-GTG-GTG-TTT-CCT-A-3’

DNA Minimum Hybridization Time with ICCD Camera

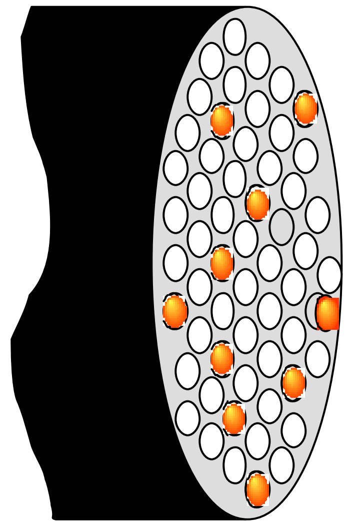
DNA Concentration	Hybridization Time (min)
1 pM	10
100 fM	20
10 fM	30
1 fM	60

Detection Limit Problem

Multiple beads provides
a signal averaging benefit.



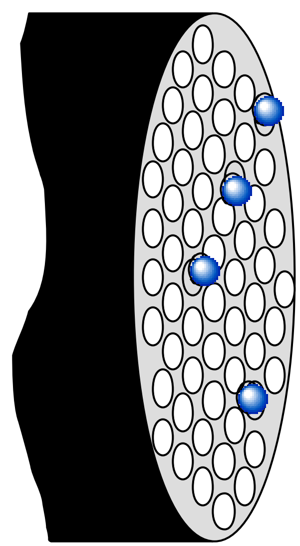
S/N increases by $\sqrt{n}$



$$\frac{\text{1000 target molecules}}{\text{10 beads}}$$

=100 target molecules/bead

Fewer beads provide more target
molecule numbers per bead.



$$\frac{\text{1000 target molecules}}{\text{4 beads}}$$

=250 target molecules/bead

Multiplexed Array Sensitivity and Selectivity with 1 fM IL2 Target Solutions

IL2 Target - 1 fM concentration - 12 hour hybridization time

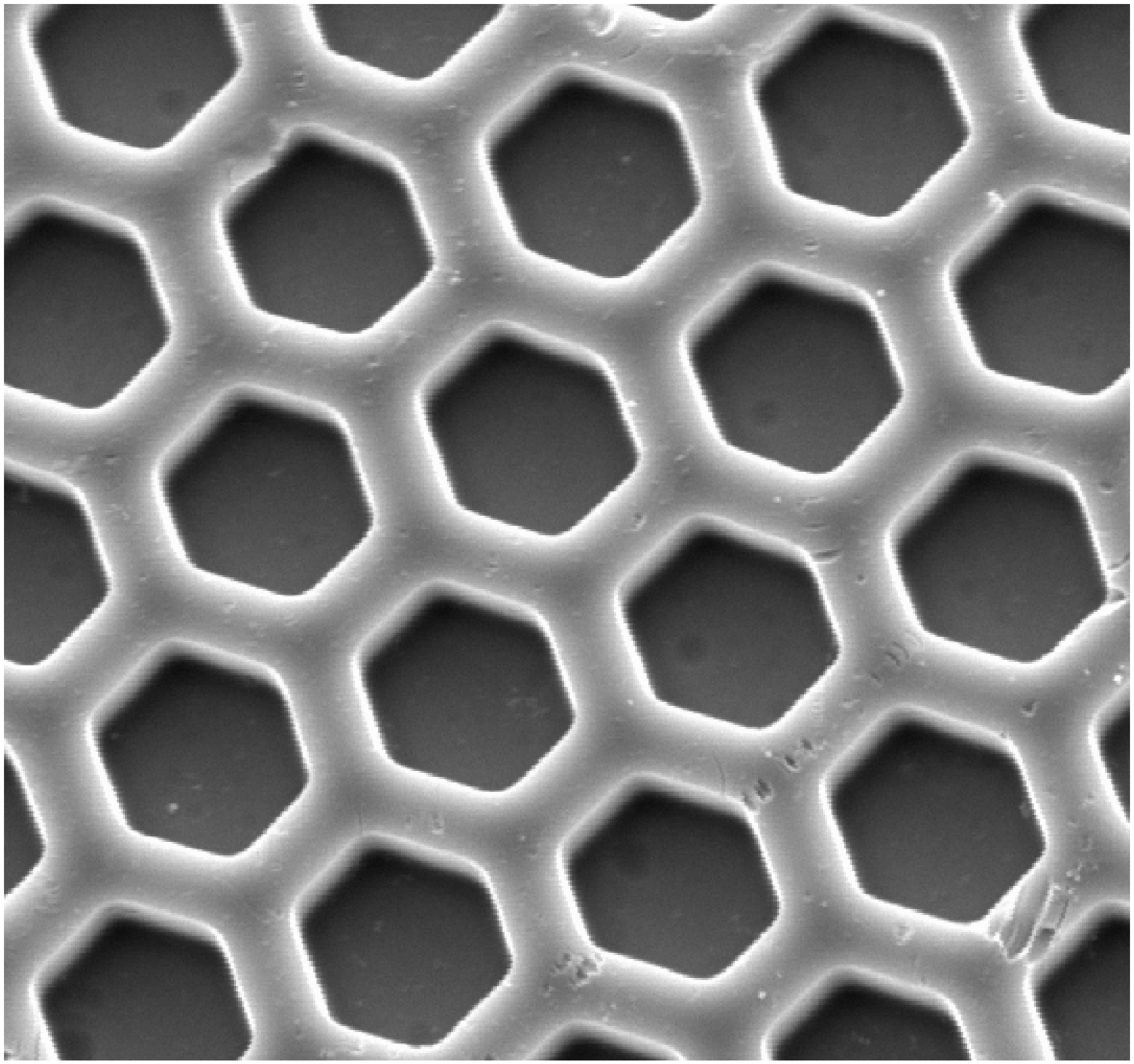
Target/Probe		Mean background ± s.d	Hybridization ± s.d.	Signal ± s.d
F508	F508C	530.43 ± 1.8	550.17 ± 7.5	{19.74} ± 7.7
	IL2	563.99 ± 7.7	677.08 ± 8.1	113.09 ± 11
	IL6	445.99 ± 3.9	449.16 ± 1.4	{3.17} ± 4.1
IL2	F508C	439.64 ± 3.5	443.34 ± 5.6	{3.70} ± 6.6
	IL2	432.52 ± 5.6	503.31 ± 6.6	70.79 ± 8.7
	IL6	431.11 ± 2.1	432.13 ± 2.8	{1.02} ± 3.5
IL6	F508C	454.84 ± 3.6	465.82 ± 1.4	{10.98} ± 3.8
	IL2	429.42 ± 0.92	517.38 ± 2.6	87.96 ± 2.8
	IL6	459.81 ± 3.0	467.82 ± 5.3	{8.01} ± 6.1

Microsphere Array Sensitivity and Selectivity with 100 aM IL2 Target Solutions

IL2 Target - 100 aM concentration - 12 hour hybridization time

Probe/Target		Mean background ± s.d	Hybridization ± s.d.	Signal ± s.d
IL2	F508C	386.97 ± 3.2	387.98 ± 1.4	{1.01} ± 3.5
	IL2	378.55 ± 2.3	394.00 ± 3.7	15.32 ± 4.4
	IL6	382.80 ± 6.3	393.81 ± 6.1	{11.01} ± 7.1
IL2	F508C	268.66 ± 2.3	274.22 ± 8.5	{5.56} ± 8.8
	IL2	297.73 ± 2.3	310.02 ± 2.3	12.29 ± 3.2
	IL6	247.59 ± 2.7	248.70 ± 6.9	{1.11} ± 7.4
IL2	F508C	410.73 ± 2.6	413.63 ± 2.6	{2.90} ± 2.9
	IL2	410.69 ± 2.7	455.26 ± 6.5	44.57 ± 7.0
	IL6	390.24 ± 7.4	392.88 ± 2.8	{2.64} ± 7.9

SEM of a Microwell Array

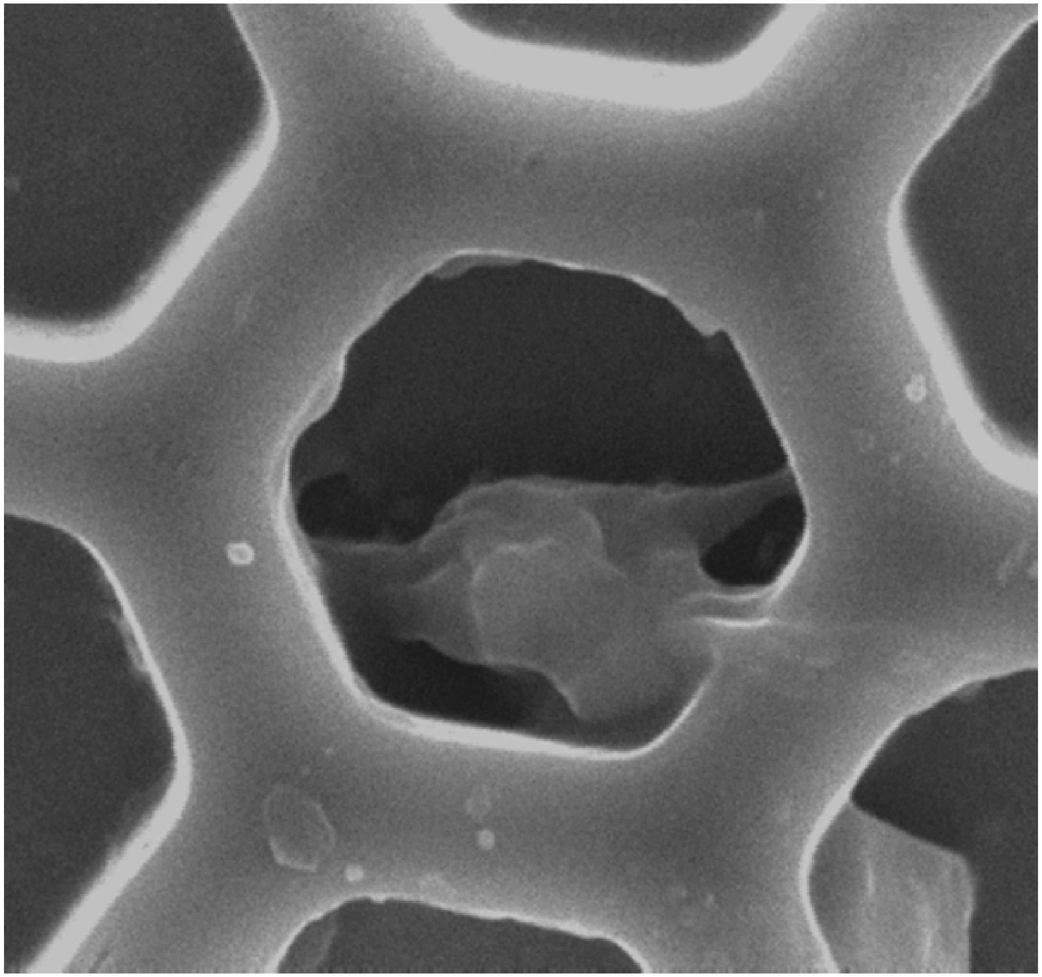


7 ? m well diameter

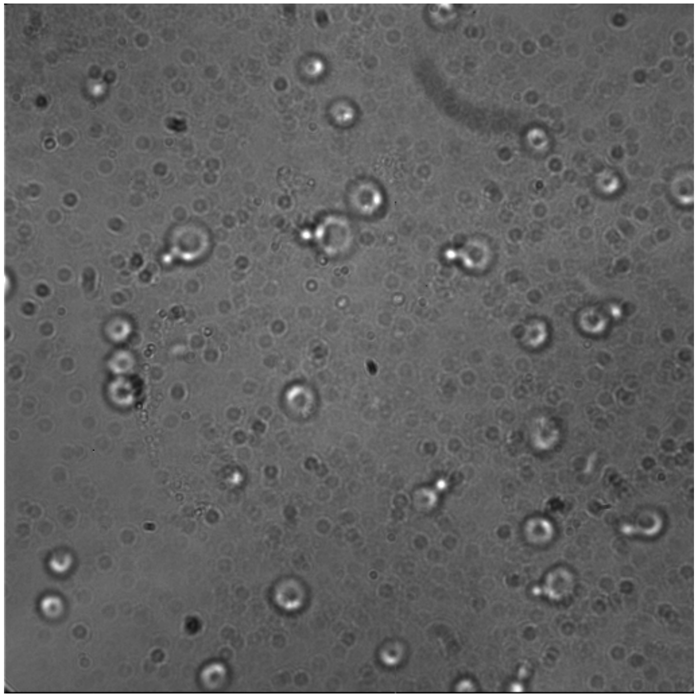
~3 ? m well depth

~90 fL well volume

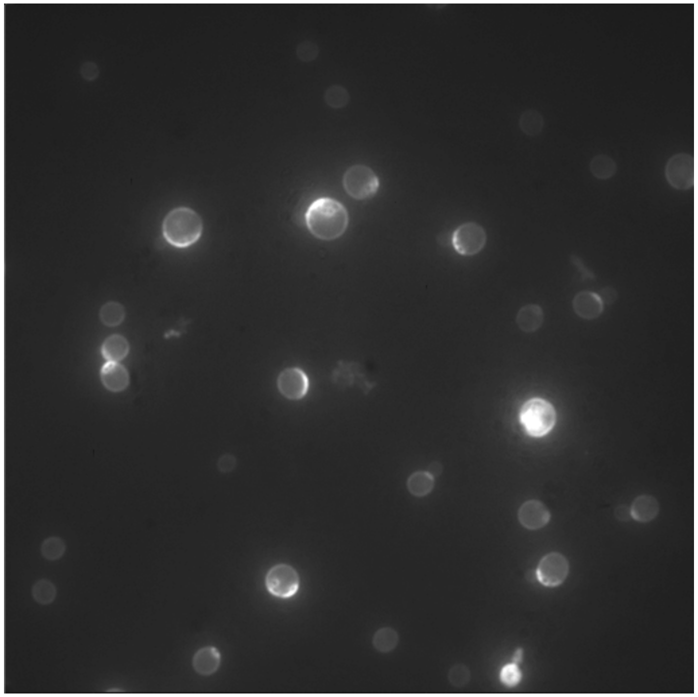
Single NIH 3T3 Mouse Fibroblast Cell in a Fiber-optic Microwell



Single Yeast (*Saccharomyces cerevisiae*) Cells Array

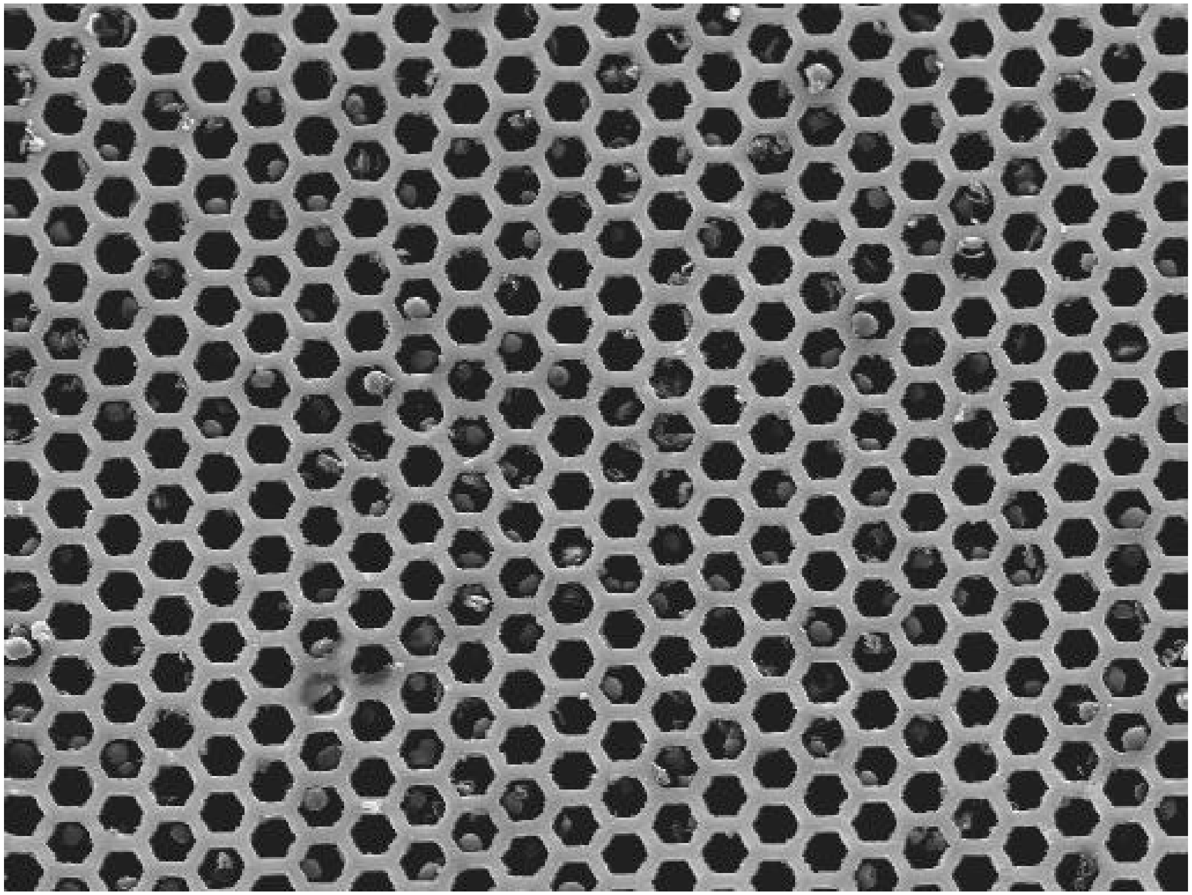


White light

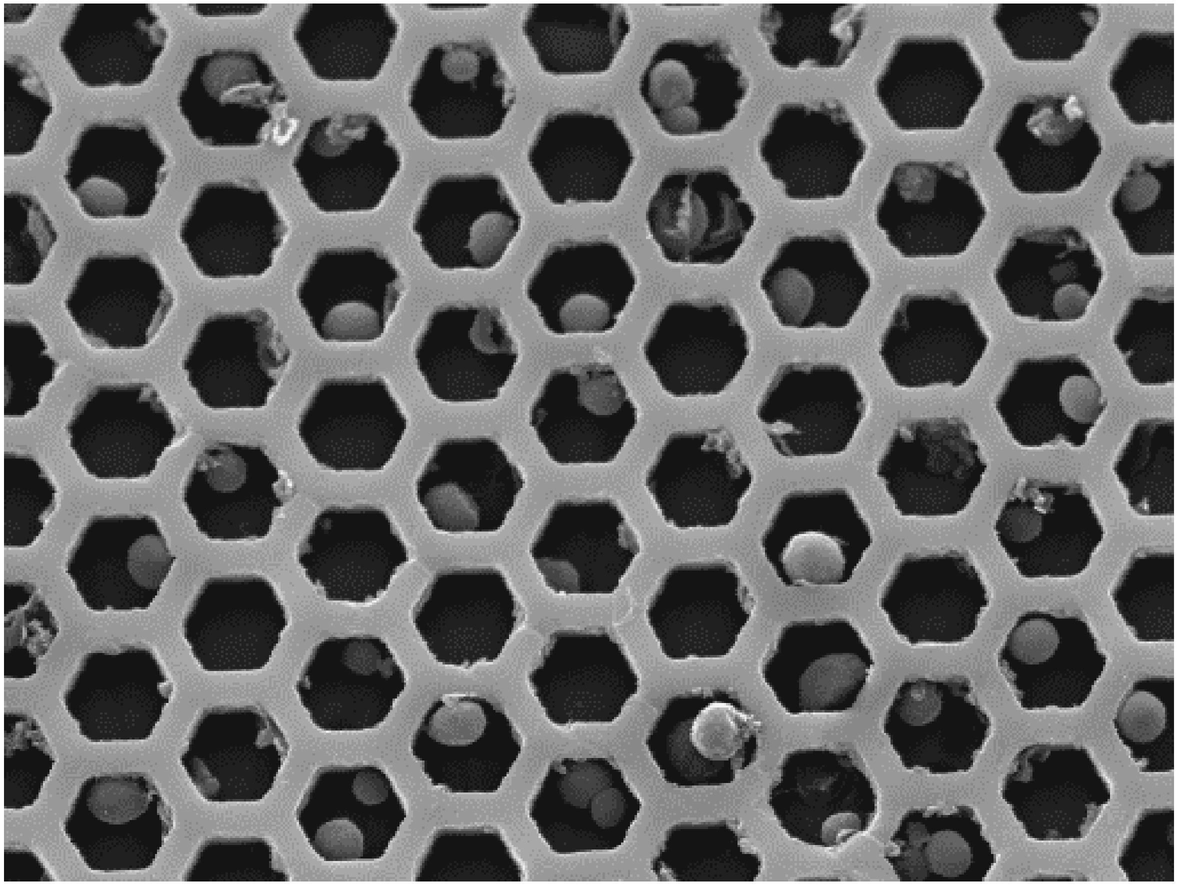


Calcofluor White 360/440

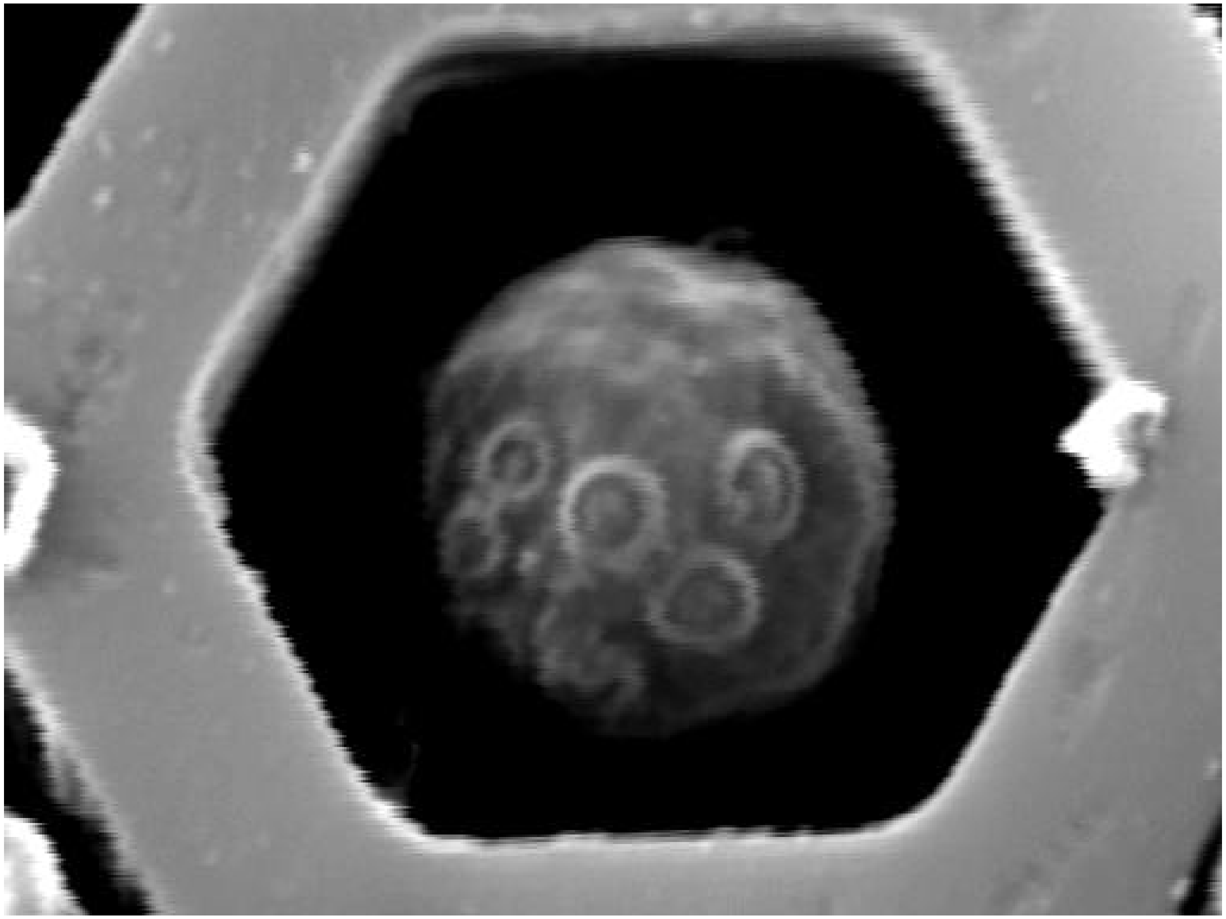
SEM images of Single Yeast Cells on the Microwells array



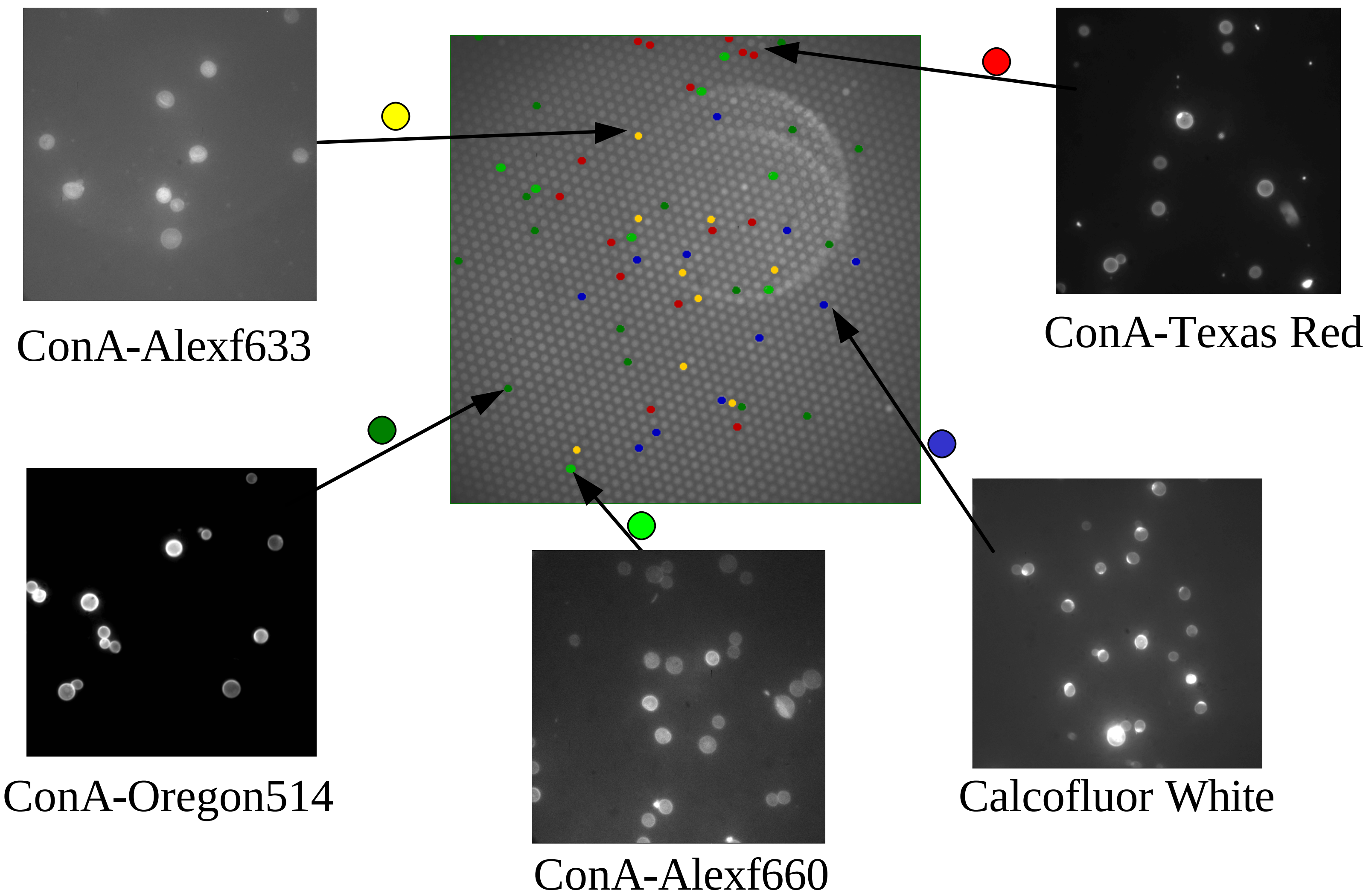
SEM images of Single Yeast Cells on the Microwells array



SEM images of Single Yeast Cells on the Microwells array



Encoded Yeast Cells on the Fiber Array

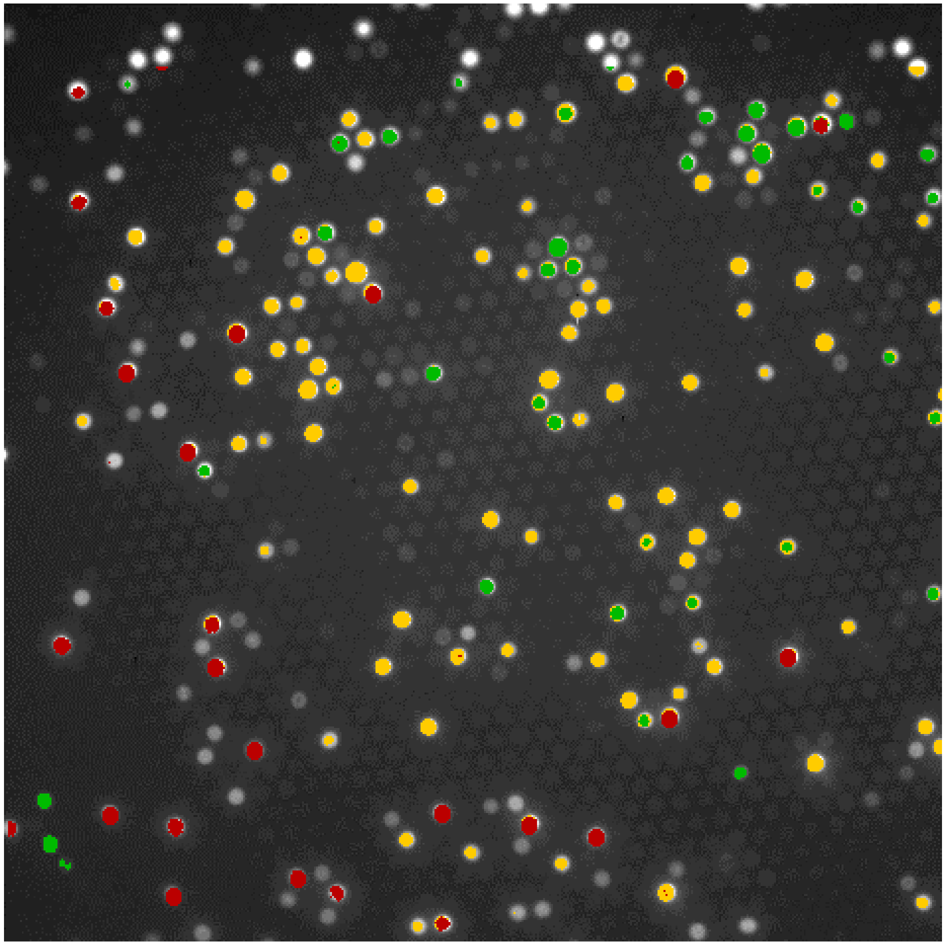


pH Measurement of Single Yeast Cells Microenvironment in the Array

Concanavalin A-FITC

Concanavalin A-FITC
+
Concanavalin A-Alexa fluor 660

Concanavalin A-FITC
+
Concanavalin A-Texas Red



Smarter Sensors- Anticipatory

Is it bad?

What does it resemble?

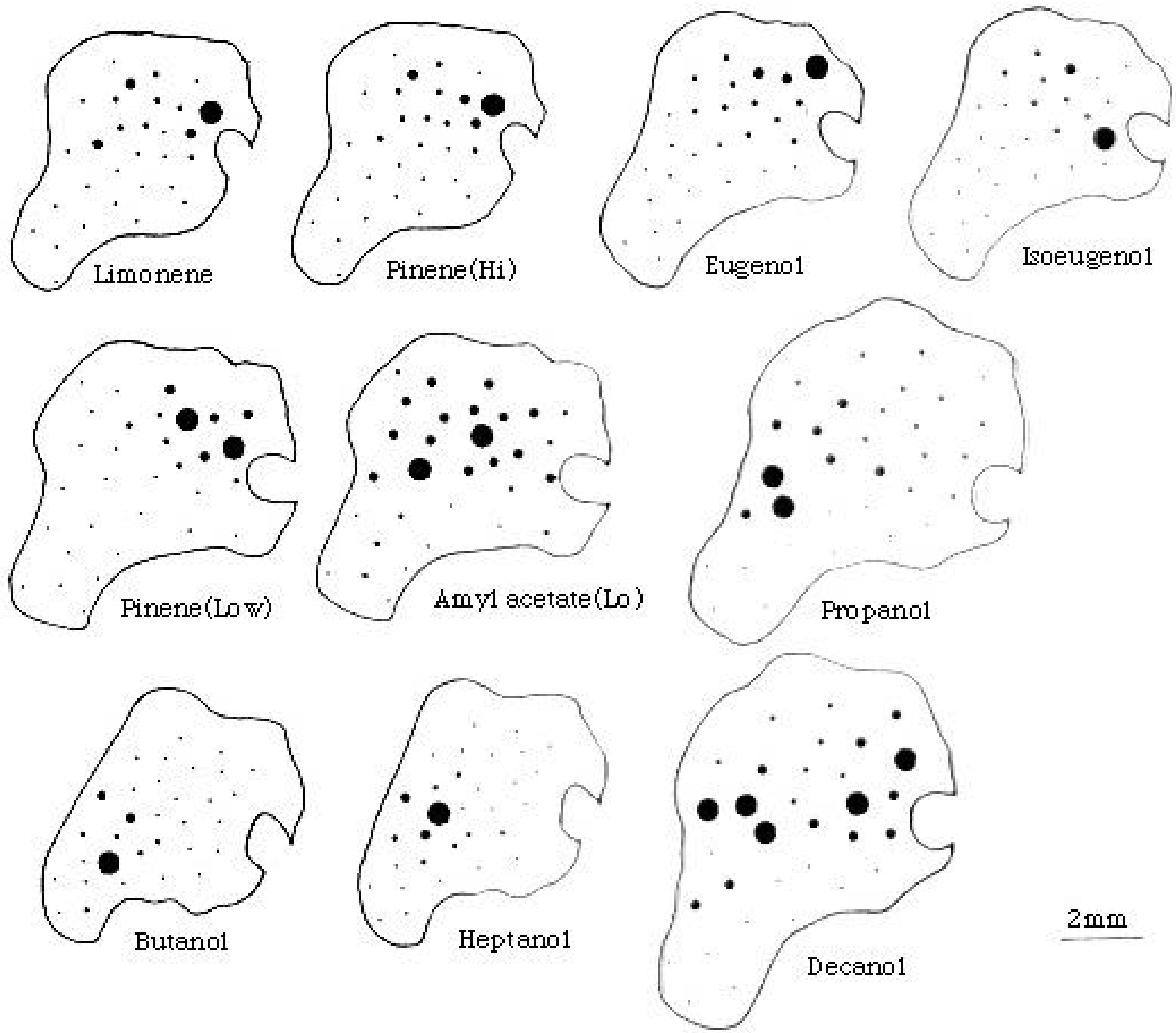
What will it do?

e.g. GI, neurotoxic, etc.

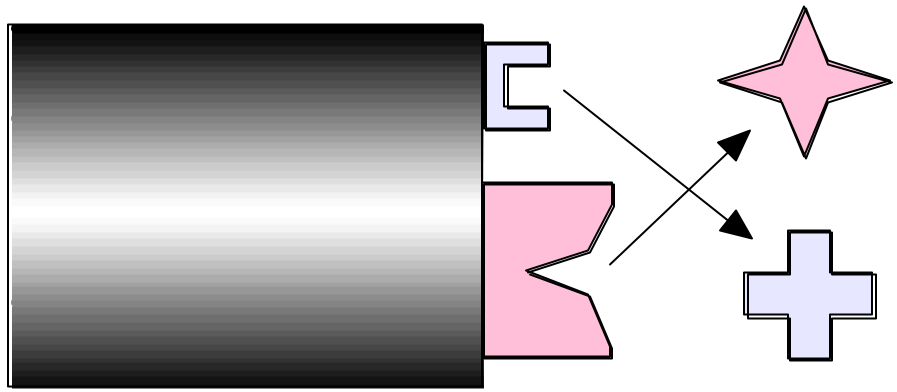
"common virulence mechanisms"

↓

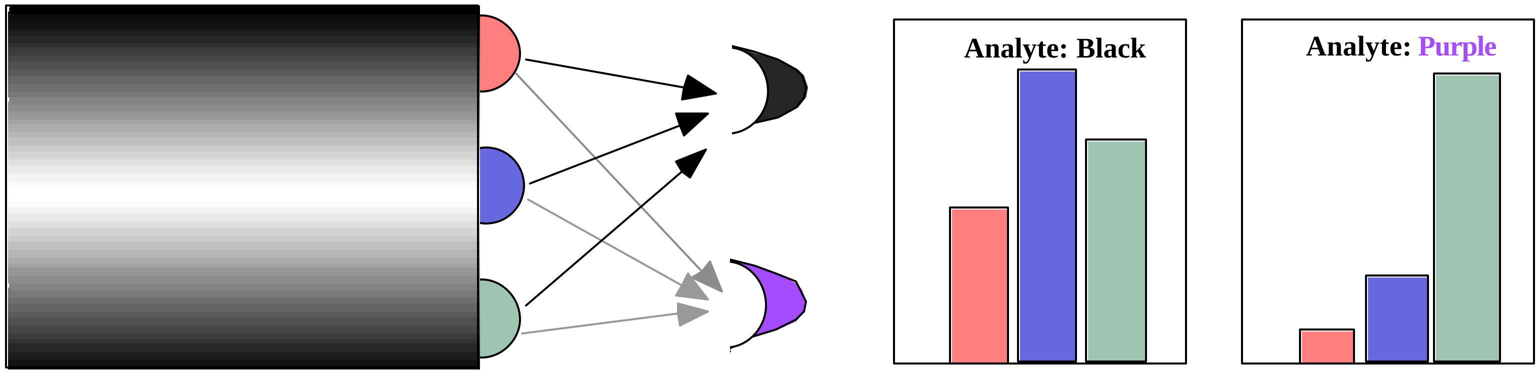
surrogates


Limonene
Pinene(Hi)
Eugenol
Isoeugenol
Pinene(Low)
Amyl acetate(Lo)
Propanol
Butanol
Heptanol
Decanol
2mm

Sensor Design

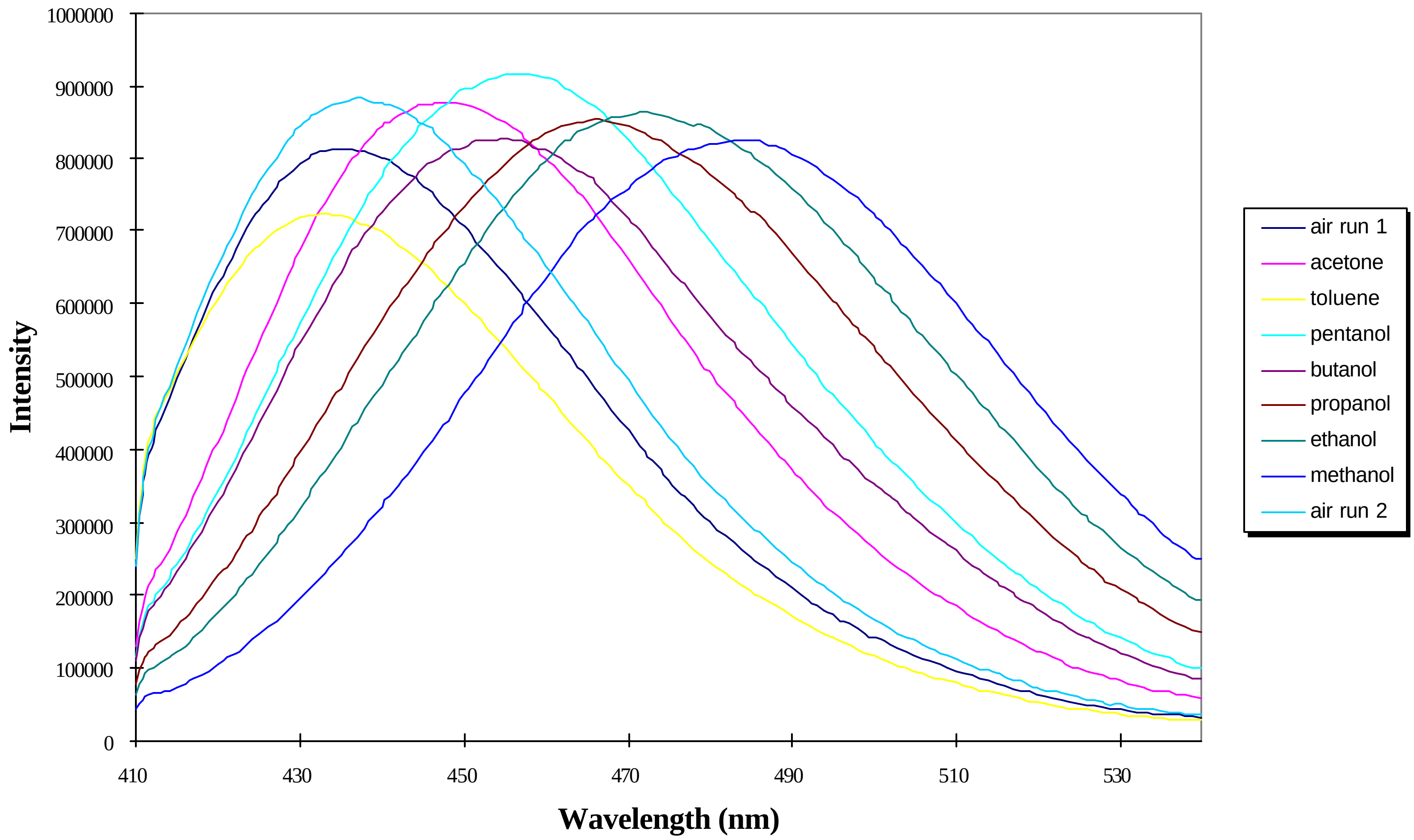


A) Lock-and-key Sensor

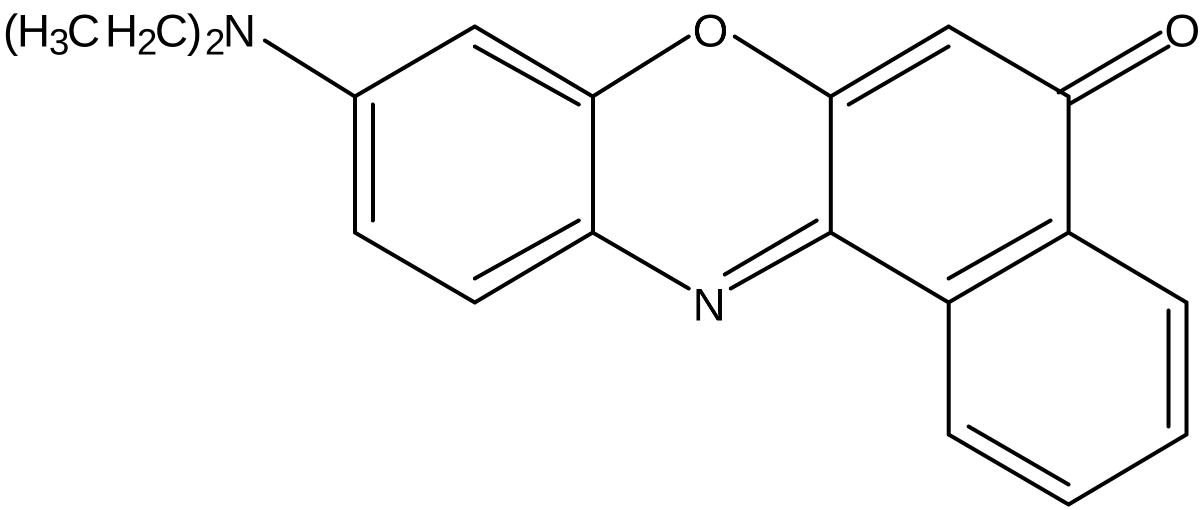


B) Cross-reactive Sensor

Solvatochromic Effect

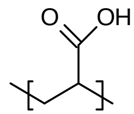


Nile Red

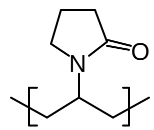


Role of Polymer Polarity

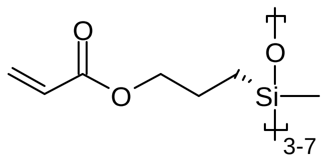
poly(acrylic acid), **PAA**



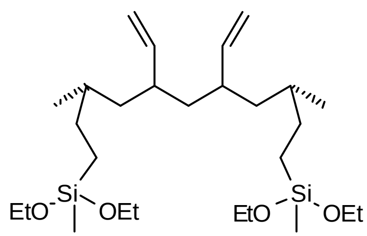
poly(N-vinyl pyrrolidone), **PVP**



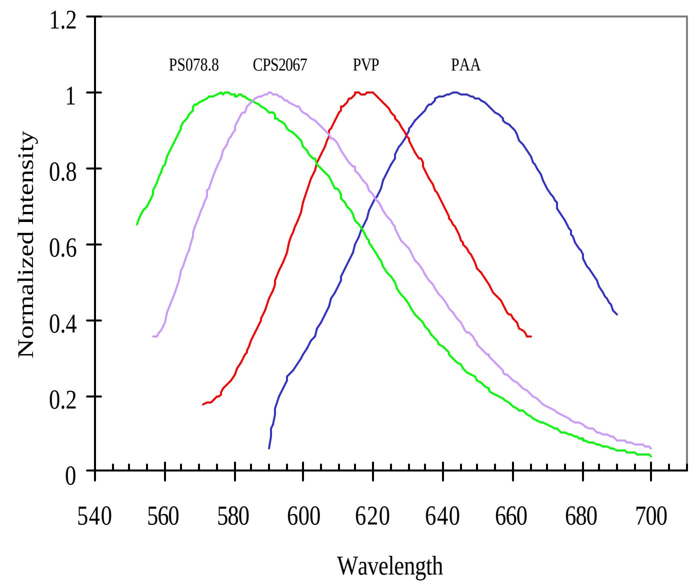
acryloxypropylmethyl-
cyclosiloxane, **CPS2067**



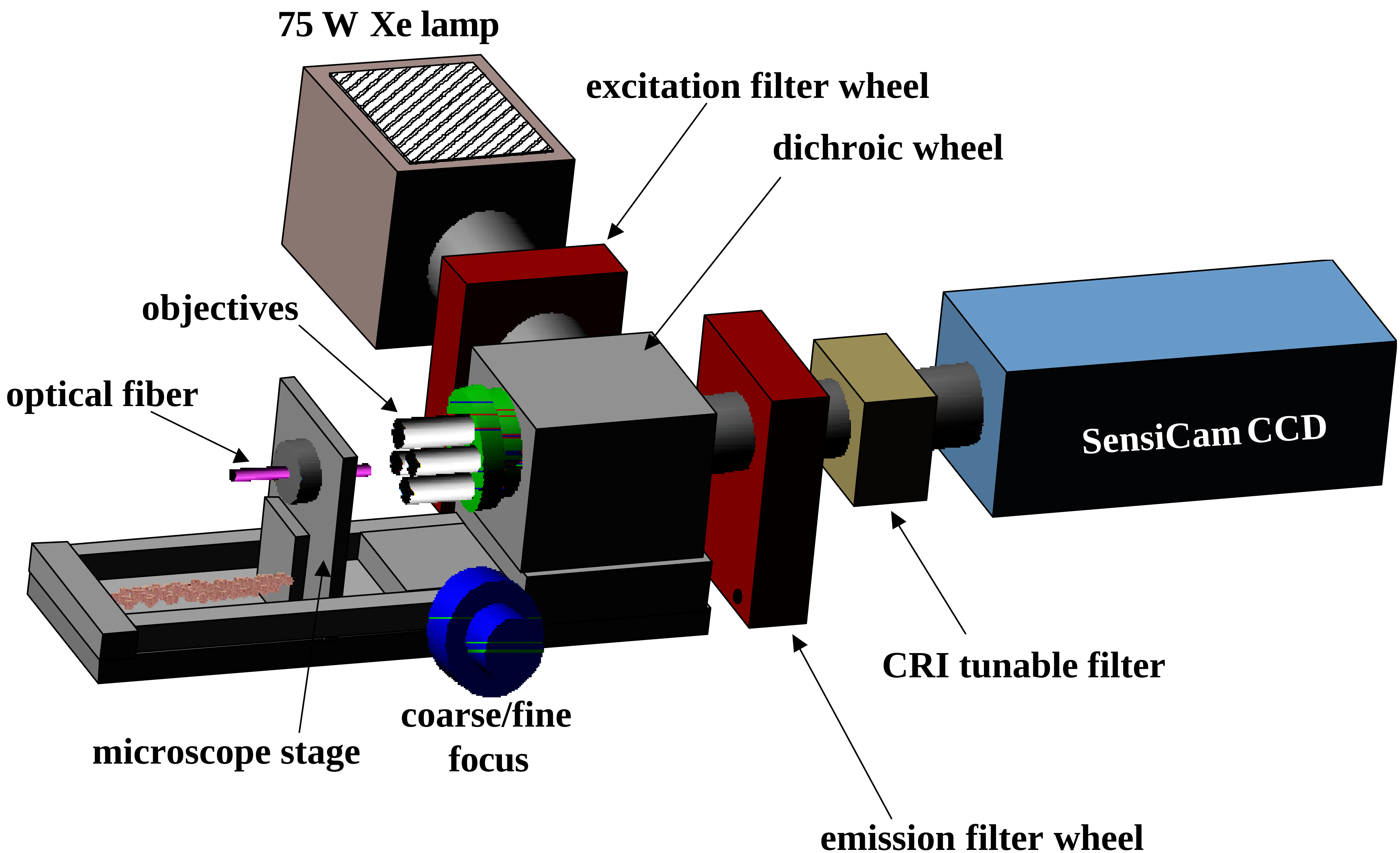
diethoxymethylsilyl-modified
polybutadiene, **PS078.8**



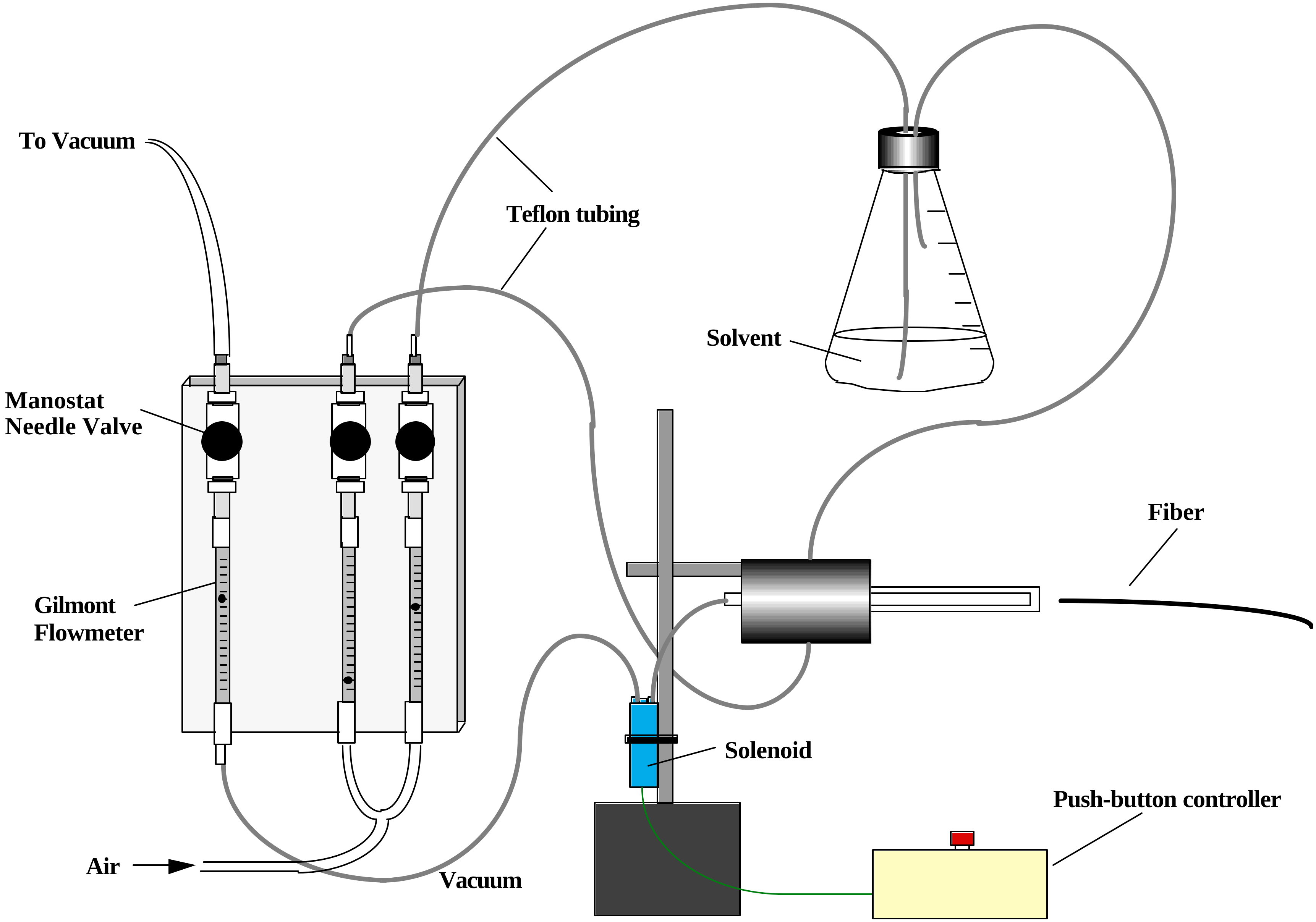
Decreasing polarity



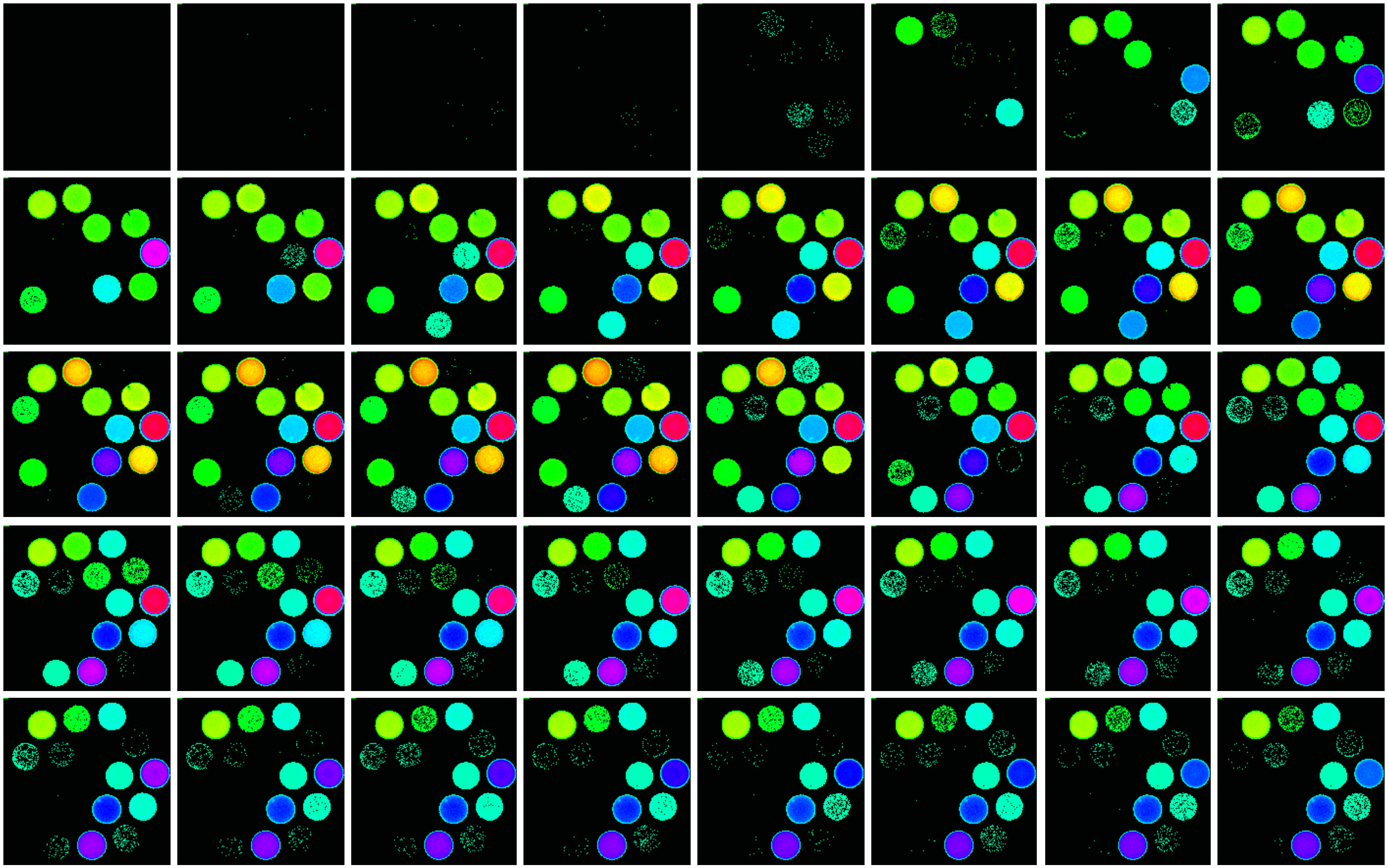
CCD-based imaging system



The Olfactometer — J. Kauer

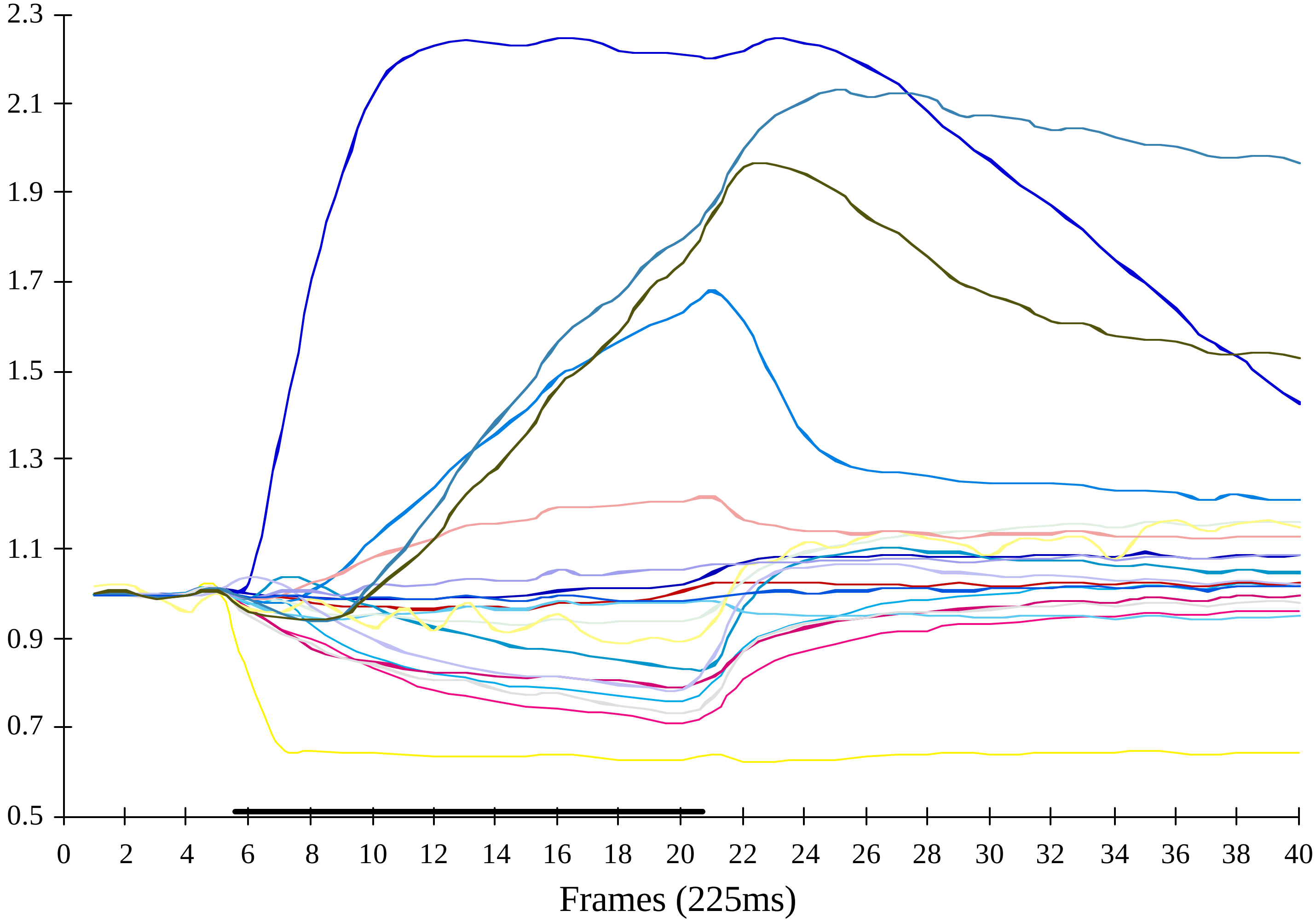


Sensor Array Response to Benzene Vapor Pulse



Dickinson, T. A., *et al.* (1996) *Nature*, 382: 697-700.

Temporal Plots from 19-Fiber Sensor Array Response to Benzene Vapor Pulse



Classification Results

Learning Vector Quantization Approach

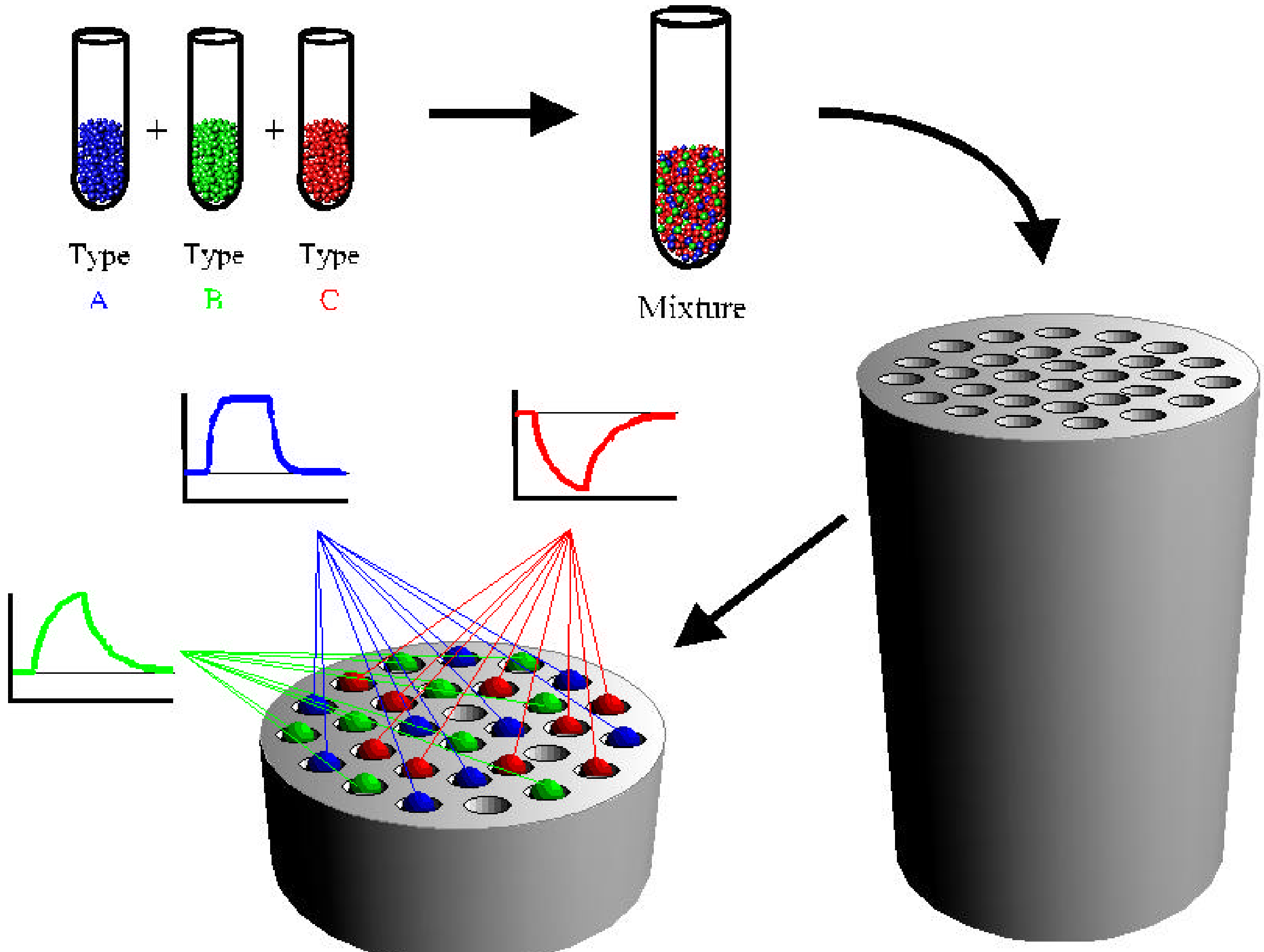
Network Output

True Identity

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	24																			
2		23										1								
3			20							3			1							
4			1	19					1					1				2		
5			1		17	4			2											
6				1	2	19					1				1					
7							18	4		1			1							
8							2	21					1							
9				1					23											
10			3							19			1			1				
11											21					1			2	
12												23						1		
13										1			23							
14									1					20			2		1	
15		1													23					
16																24				
17																	24			
18																		24		
19											4								20	
20		1																		23

1) Acetone
2) Butyl Acetate
3) Beauty
4) Camphor
5) Carvone -
6) Carvone +
7) Chloroform
8) Dichloroethane
9) DMSO
10) Drakkar Noir
11) Water
12) Heptane
13) Isopropanol
14) Indole
15) Mercaptoethanol
16) Methanol
17) Propanol
18) Propionic Acid
19) Pseudoexplosive
20) Toluene

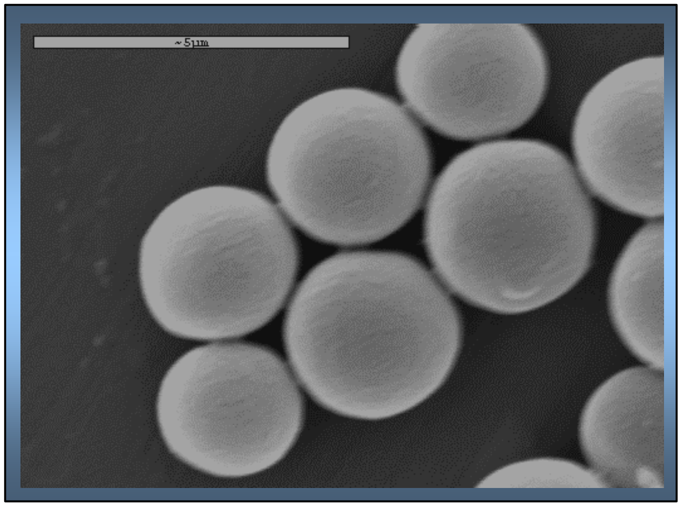
SENSOR ARRAYS are Assembled 'Randomly' in ONE Fabrication Step



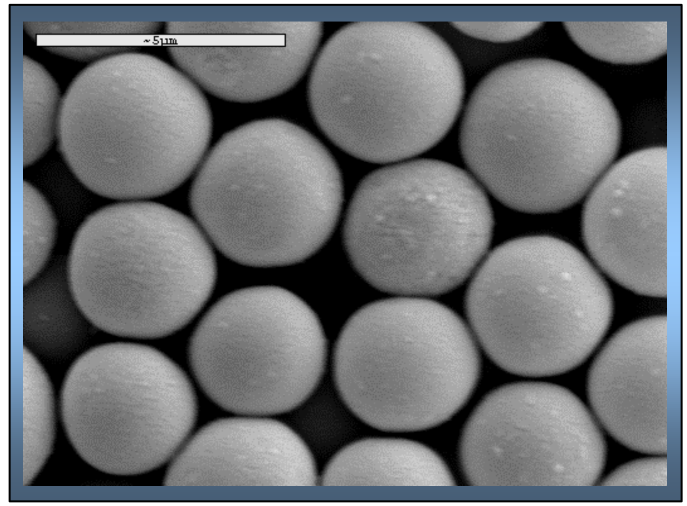
NOTE: the Sensor Array is a 'Self-Encoding' Bead Array (SEBA). Billions of Sensors are Fabricated at Once.

Dickinson, T.A. *et al.* (1999) *Anal. Chem.* 71 (11): 2192-2198.

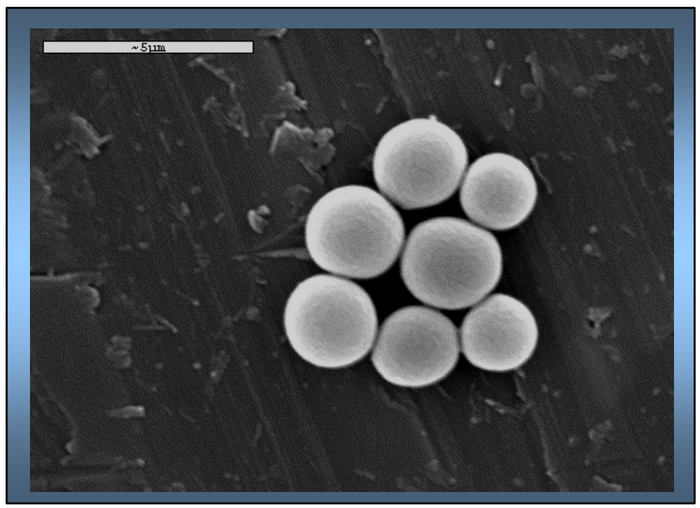
Hollow Poly(benzyl methacrylate) Spheres



3.5 h polymerization



6.5 h polymerization

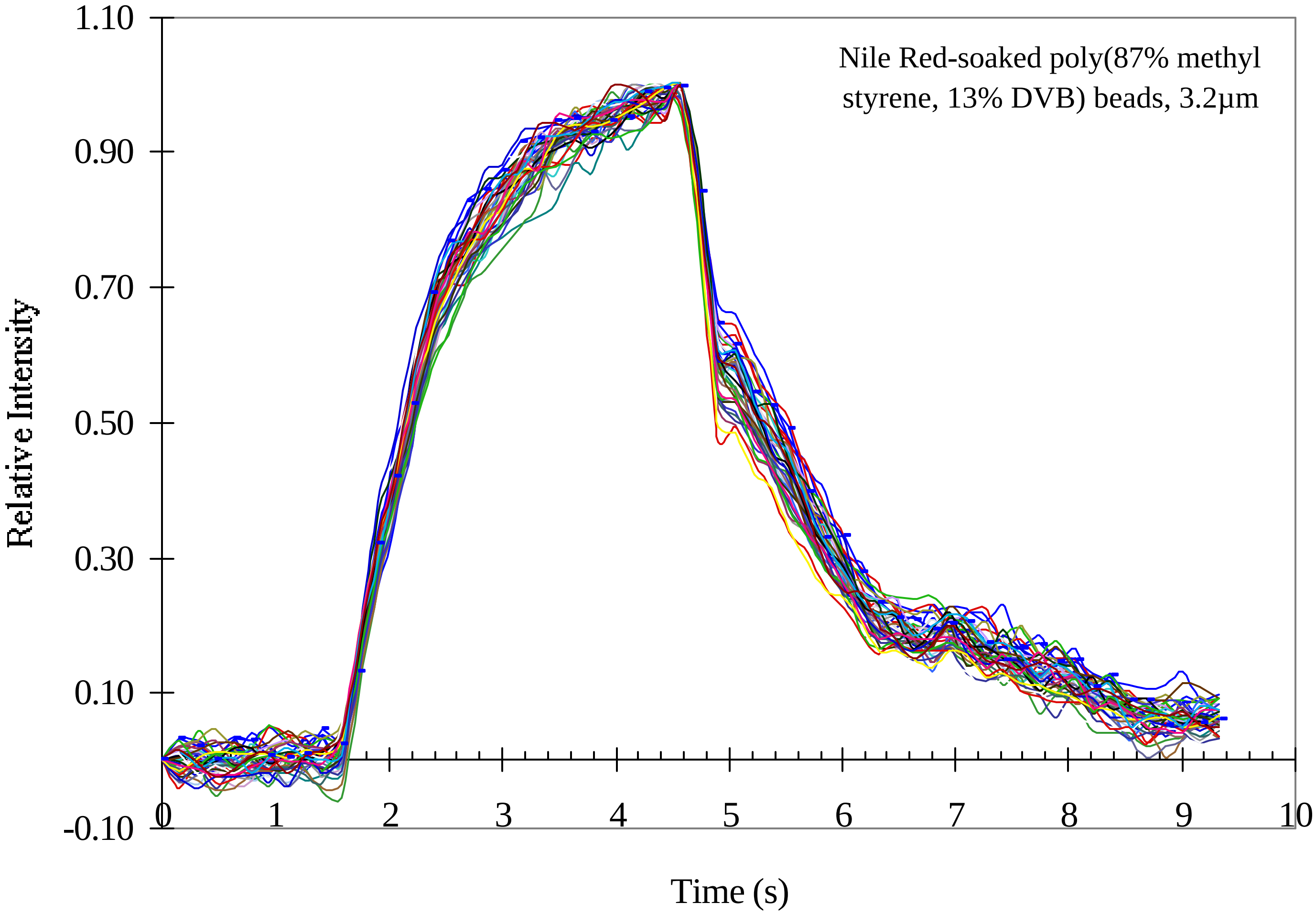


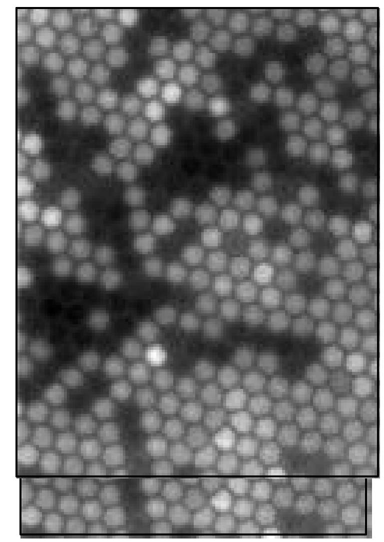
14h polymerization

Chem. Mater. 2000

Nile Red/PolyMethylStyrene Beads in Wells:

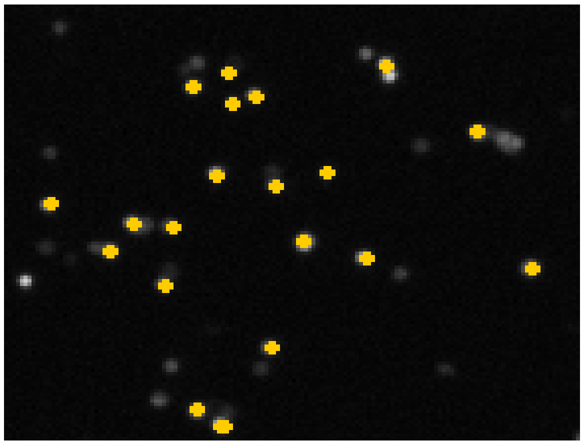
Response of 40 beads to methanol pulse





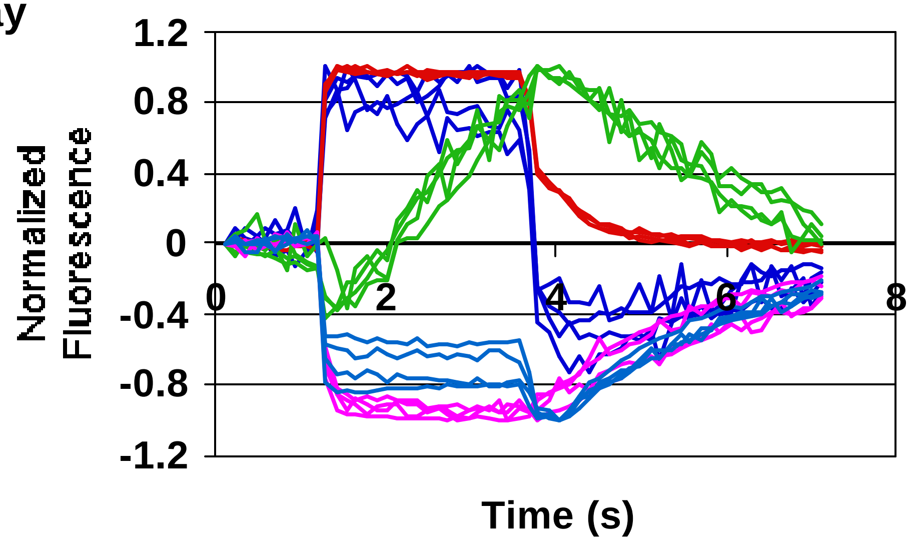
bright = wells with beads
dark = empty wells

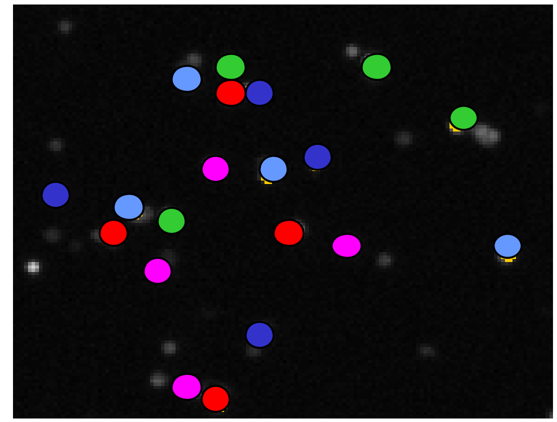
Sensor Registration Problem



Randomly Ordered
5 Bead Array

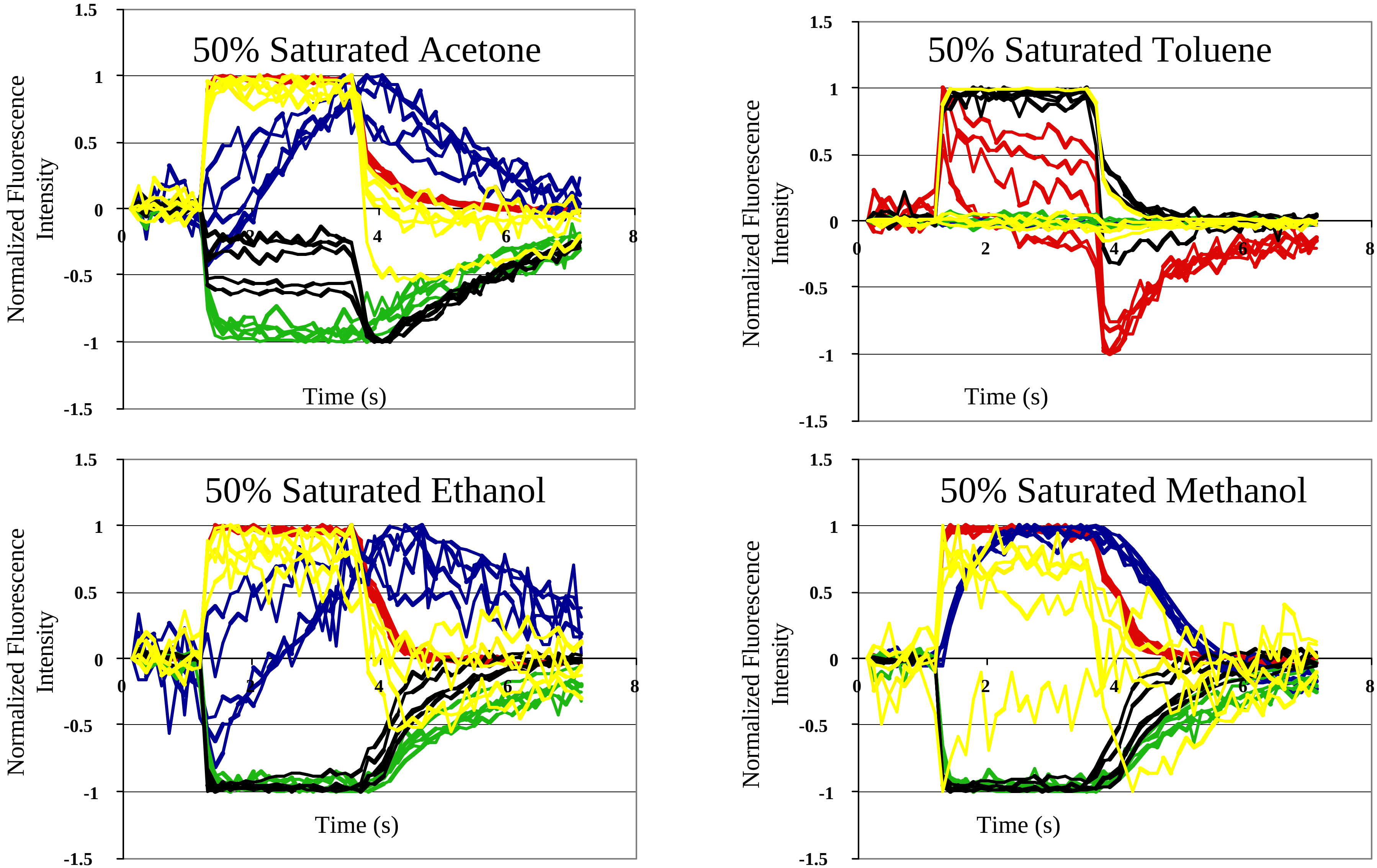
Response of Bead Array to Methanol

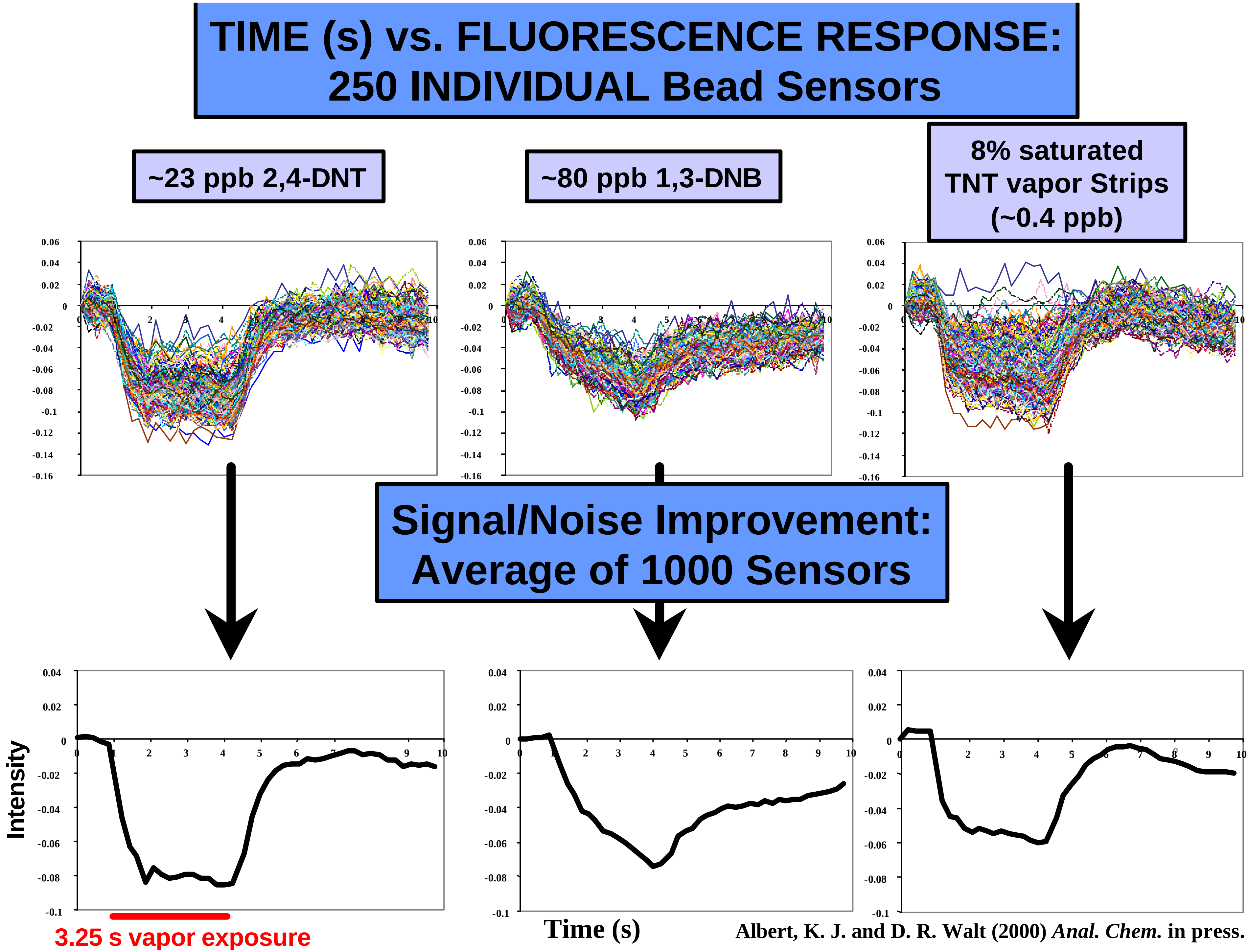




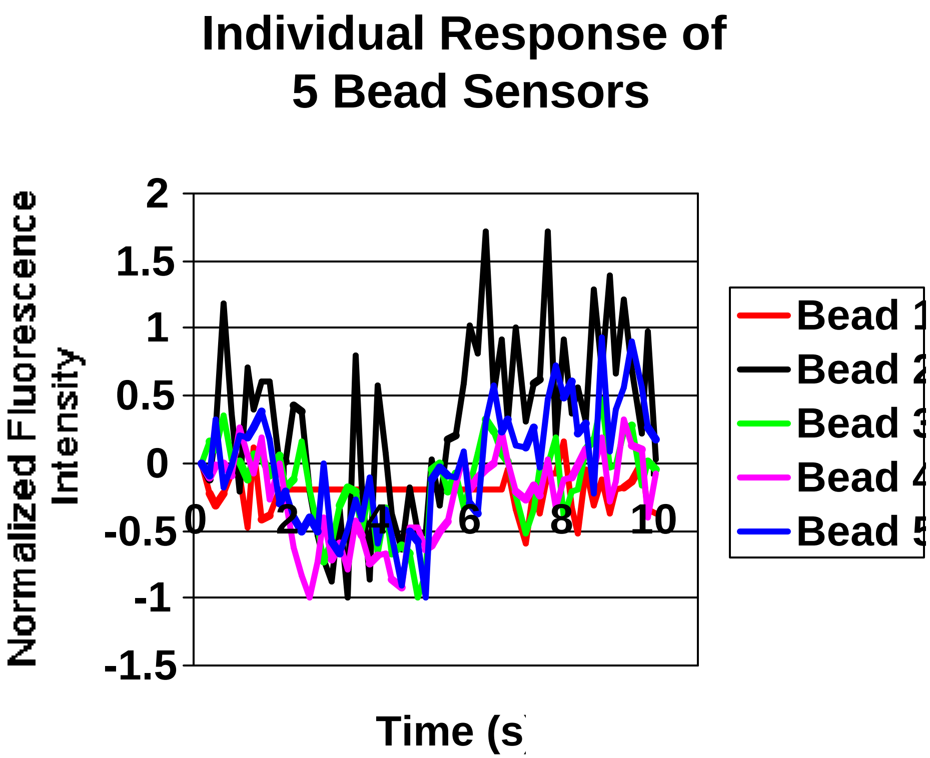
Decoded 5
Bead Array

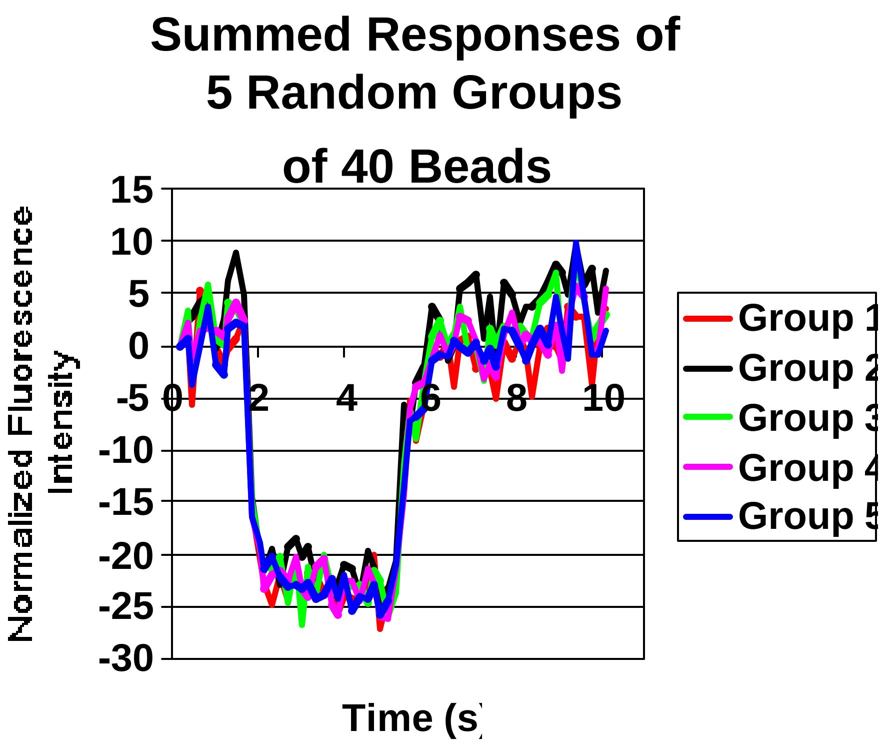
5 Sensor Types with 4 Analytes




TIME (s) vs. FLUORESCENCE RESPONSE:
250 INDIVIDUAL Bead Sensors
~23 ppb 2,4-DNT
~80 ppb 1,3-DNB
8% saturated
TNT vapor Strips
(~0.4 ppb)
Signal/Noise Improvement:
Average of 1000 Sensors
Intensity
3.25 s vapor exposure
Time (s)
Albert, K. J. and D. R. Walt (2000) *Anal. Chem.* in press.

Signal Summing





Summing improves signal-to-noise ratio.

Analytes for Two Class Problem

- Pure Analytes
  - Acetone
  - Benzene
  - Chloroform
  - Ethanol
  - Ethyl Acetate
  - Heptane
  - Methanol
  - Toluene
  - 1,3-Dinitrobenzene
  - 4-Nitrotoluene
- Binary Mixtures
  - Ethyl Acetate/Heptane
  - Methanol/Benzene
  - 4-NT/Benzene
  - 4-NT/Heptane
  - 4-NT/Methanol
  - 1,3-DNB/Ethyl Acetate
  - 1,3-DNB/Heptane

Concentrations of Analytes

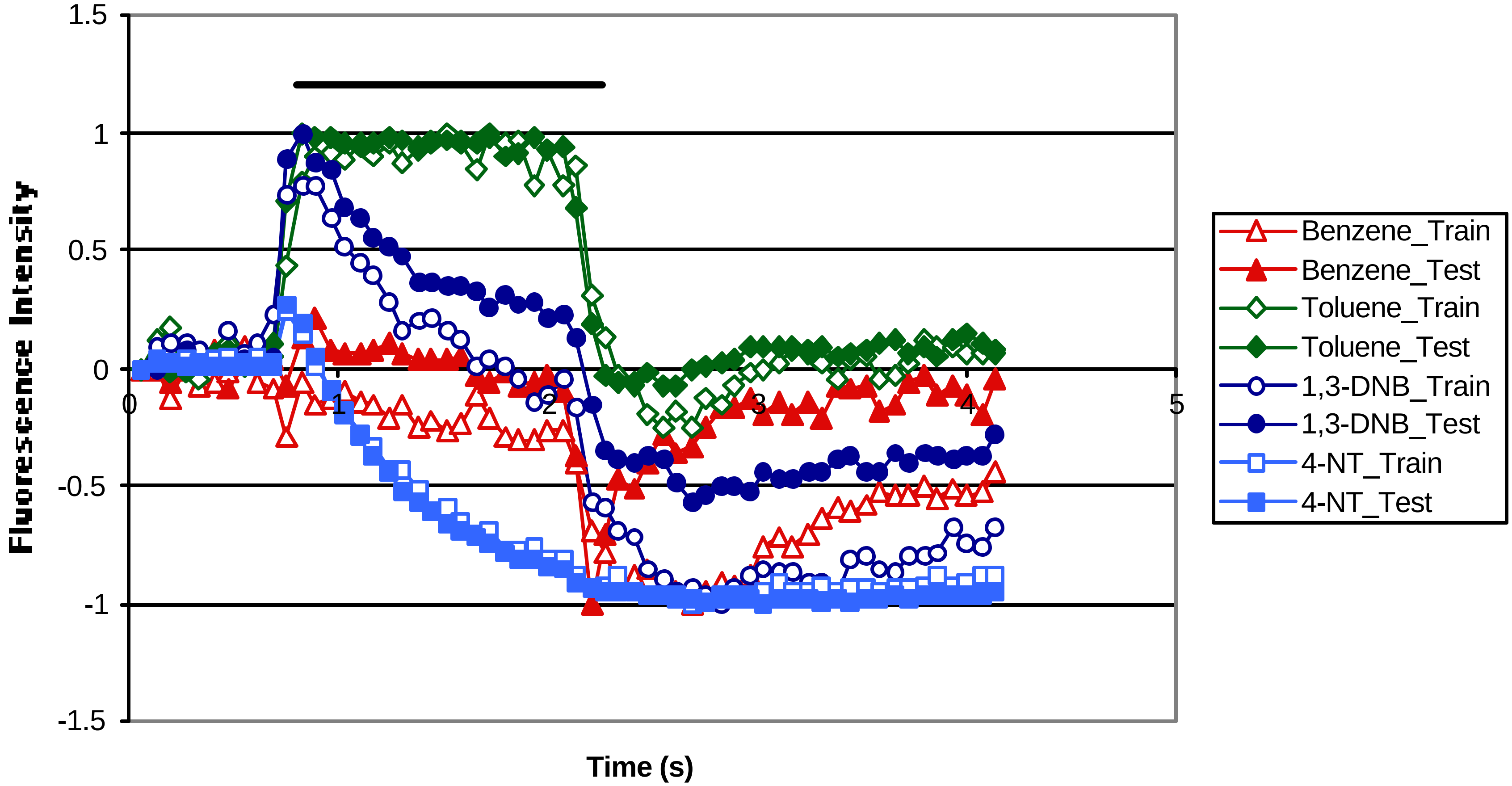
Table 1: The concentration of the pure analytes ±15%. The concentrations were calculated based on literature values for analyte vapor pressures.

Analyte	Vapor Pressure @25 °C (mmHg)	Concentration (ppm)
Acetone	2.31E+02	7.6E+04
Benzene	9.53E+01	3.1E+04
Chloroform	1.97E+02	6.5E+04
Ethanol	5.90E+01	1.9E+04
Ethyl Acetate	9.45E+01	3.1E+04
Heptane	4.57E+01	1.5E+04
Methanol	1.27E+02	4.2E+04
Toluene	2.84E+01	9.4E+03
1,3-Dinitrobenzene	9.00E-04	6.0E-01
4-Nitrotoluene	1.64E-01	1.1E+02

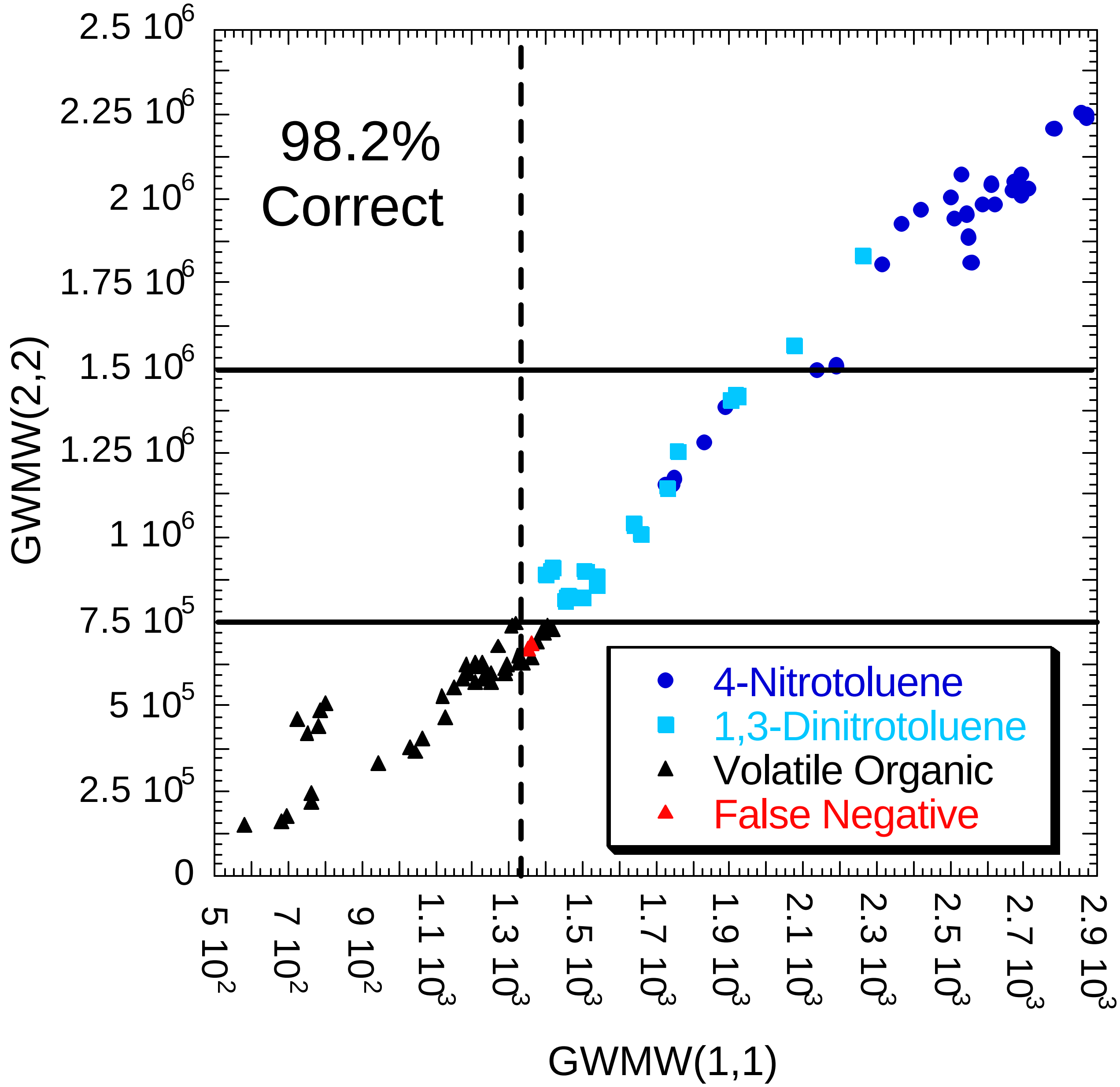
Table 2: The concentration of the binary mixtures ±15%.

Analyte 1	Analyte 2	Concentration analyte1 (ppm)	Concentration analyte2 (ppm)
Benzene	Methanol	3.1E+04	4.2E+04
Benzene	4-Nitrotoluene	3.1E+04	5.5E+01
Benzene	4-Nitrotoluene	3.1E+04	1.1E+02
Ethyl Acetate	Heptane	3.1E+04	1.5E+04
Ethyl Acetate	1,3-Dinitrotoluene	3.1E+04	3.0E-01
Ethyl Acetate	1,3-Dinitrotoluene	3.1E+04	6.0E-01
Heptane	1,3-Dinitrotoluene	1.5E+04	6.0E-01
Heptane	4-Nitrotoluene	1.5E+04	1.1E+02
Methanol	4-Nitrotoluene	4.2E+04	5.5E+01
Methanol	4-Nitrotoluene	4.2E+04	1.1E+02

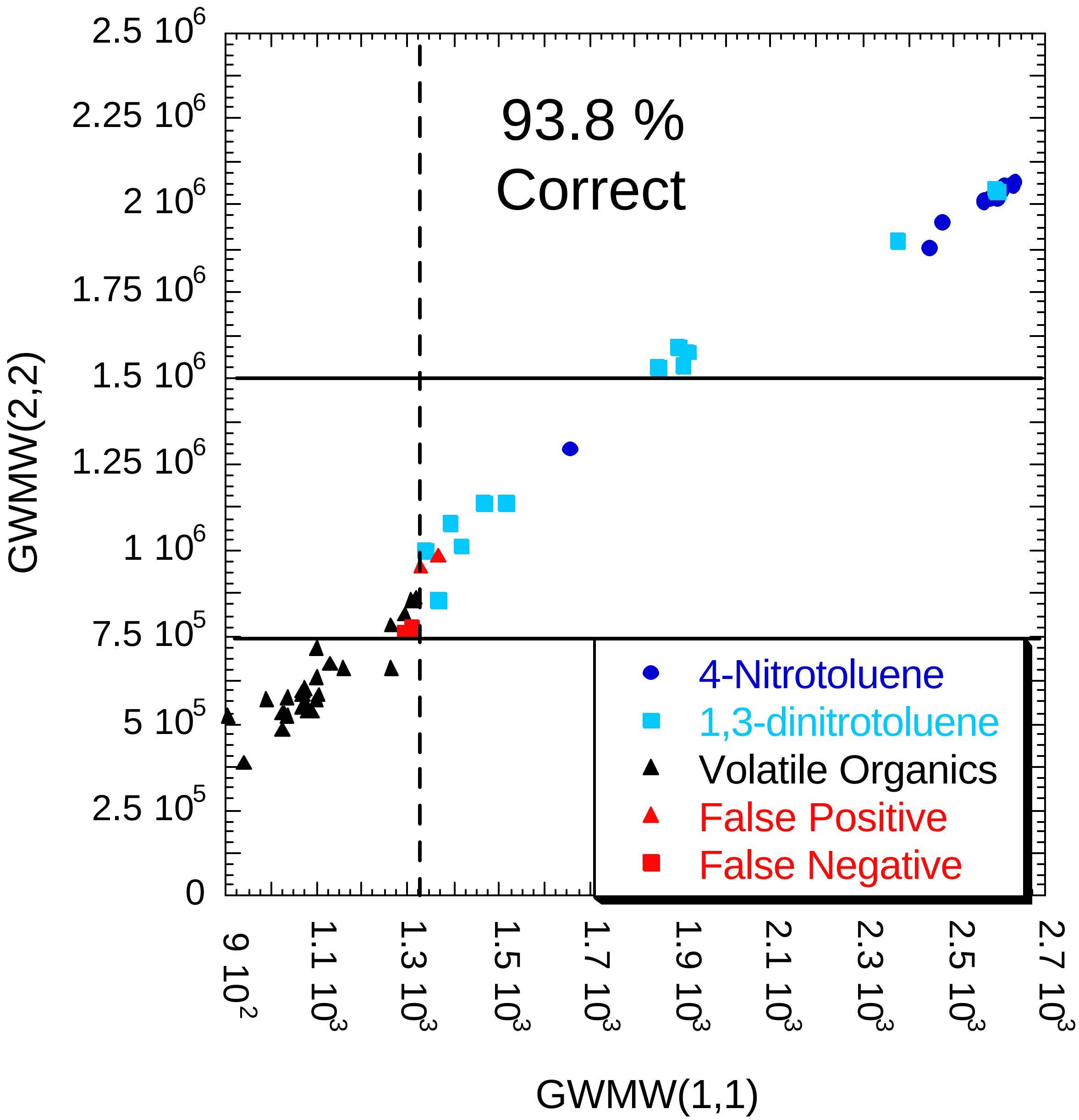
Reproducible Responses from Training to Testing array



First Testing Array (1 Month)



Second Test Array (7 months)



Live/Dead Bacteria Discrimination

Calculated Identity

Actual Identity

	Live B10	Live B4	Live B5	Live B8	Live B9	Dead B10	Dead B4	Dead B5	Dead B8	Dead B9	Medium
Live B10	4	1	0	0	0	0	0	0	0	0	0
Live B4	2	3	0	0	0	0	0	0	0	0	0
Live B5	0	0	5	0	0	0	0	0	0	0	0
Live B8	0	0	0	4	0	0	1	0	0	0	0
Live B9	0	0	0	0	5	0	0	0	0	0	0
Dead B10	0	0	0	0	0	5	0	0	0	0	0
Dead B4	0	0	0	0	0	0	5	0	0	0	0
Dead B5	0	0	0	0	0	0	0	5	0	0	0
Dead B8	0	0	0	0	0	0	0	0	5	0	0
Dead B9	0	0	1	0	2	0	0	2	0	0	0
Medium	0	0	0	0	0	0	0	0	0	0	10

85% Correct, 87% Variance (7PCs)

B10: *Acintobacterium*
B4: *M. luteus*
B5: *E. coli*
B8: *Salmonella*
B9: *Klebsiella peumoniae*

Acknowledgements

ONR

DARPA

DOE

NIH

NSF

Walt Group

Myoyong Lee
Karri Michael
Paul Pantano
Caroline Schauer
Jenny Tam

Keith Albert
Todd Dickinson
Jane Ferguson
Michael Fleming
Jason Epstein

Ferenc Szurdoki
Frank Steemers
Shannon Stitzel
Laura Taylor
Tarun Mandal

Lawrence Livermore

Fred Milanovich

Pennsylvania State University

Peter C. Jurs

Tufts University School of Medicine

John S. Kauer   Joel White

Johns Hopkins University

Lenore Cowen

Illumina

Mark Chee   Kevin Gunderson